

**TRANSMITTAL LETTER TO THE UNITED STATES
DESIGNATED/ELECTED OFFICE (DO/EO/US)
CONCERNING A FILING UNDER 35 U.S.C.371****UMV-1474US**

U.S. APPLICATION NO. (if known, see 37 CFR 1.5)

09/600766

INTERNATIONAL APPLICATION

PCT/US99/01382

INTERNATIONAL FILING DATE

21 January 1999 (21.01.99)

PRIORITY DATE CLAIMED

21 January 1998 (21.01.98)

TITLE OF INVENTION

**TARGETING GENE TRANSFER VECTORS TO CERTAIN CELL TYPES BY PSEUDOTYPING
WITH VIRAL GLYCOPROTEIN**

APPLICANT(S) FOR DO/EO/US

Gary J. NABEL; Rafael DELGADO; and Zhi-yong YANG

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C.371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This express request to begin national examination procedures (35 U.S.C. 371 (f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371 (b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
 - a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☒ has been transmitted by the International Bureau.
 - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
 - a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☐ have been transmitted by the International Bureau.
 - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
 - d. ☒ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)) **(unexecuted) (4 sheets);**
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included
13. ☐ A FIRST preliminary amendment.
☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information: **Transmittal Letter (2 sheets in duplicate); Response to the Invitation to Correct Defects in the International Application with Sequence Listing attached along with Sequence listing diskette (34 sheets); PCT International Published Application (WO 99/37331) (with International Search Report); (76 sheets); Check (\$840); Certificate of Express Mailing (1 sheet); and Return Postcard.**

U.S. APPLICATION NO. (if known, see 37 CFR 1.5)

INT. APPLICATION NO.

ATTORNEY'S DOCKET NO.

09/600766

PC/US99/01382

UMV-1474US

17. ☒ The following fees are submitted:**BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) (a/o January 1, 2000):**

Search Report has been prepared by the EPO or JPO.....\$970
 International preliminary examination fee paid to
 USPTO (37 CFR 1.482).....\$840
 No international preliminary examination fee paid to
 USPTO (37 CFR 1.482) but international search fee
 paid to USPTO (37 CFR 1.445(a)(2)).....\$690
 Neither international preliminary examination fee
 (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2))
 paid to USPTO.....\$670
 International preliminary examination fee paid to
 USPTO (37 CFR 1.482) and all claims satisfied provisions
 of PCT Article 33(2)-(4).....\$96

ENTER APPROPRIATE BASIC FEE AMOUNT =**CALCULATIONS PTO USE ONLY****\$840**

Surcharge of **\$130.00** for furnishing the oath or declaration later than 20 30
 months from the earliest claimed priority date (37 CFR 1.492(e)).

\$--

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE
Total claims	20 -20 =	0	X \$18.00
Independent claims	2 -3 =	0	X \$78.00
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ 260.00

\$0**\$0**

\$--

TOTAL OF ABOVE CALCULATIONS =**\$840**

Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity
 Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28)

\$--

SUBTOTAL =**\$840**

Processing fee of **\$130.00** for furnishing the English translation later than ☐ 20 ☐ 30
 months from the earliest claimed priority date (37 CFR 1.492(f)).

\$--

TOTAL NATIONAL FEE =**\$840**

Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment
 must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31).
\$40.00 per property

\$--

TOTAL FEES ENCLOSED =**\$840**Amount to be:
refunded

\$

charged

\$

a. ☒ A check in the amount of **\$840** to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. _____ in the amount of \$ _____ to cover the above fees.
 A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit
 any overpayment to Deposit Account No. **12-0080**. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

DeAnn F. Smith
LAHIVE & COCKFIELD, LLP
28 State Street
Boston, Massachusetts 02109
United States of America
(617)227-7400
Date: 21 July 2000

SIGNATURE

DeAnn F. Smith

NAME

36,683

REGISTRATION NUMBER

534 Rec'd PCT/PTC 21 JUL 2000

**TARGETING GENE TRANSFER VECTORS TO
CERTAIN CELL TYPES BY PSEUDOTYPING WITH
VIRAL GLYCOPROTEIN**

FIELD OF THE INVENTION

5 The present invention relates generally to compositions and methods for selective gene transfer, and in particular, to methods for targeting genes to certain cell types, comprising introducing to a cell population the gene to be transferred operatively-linked to an appropriate transfer vehicle, wherein the transfer vehicle is associated with a transmembrane form of viral glycoprotein.

10 **BACKGROUND OF THE INVENTION**

 Ebola virus has been identified as the cause of several highly lethal outbreaks of hemorrhagic fever. Infection begins typically with flu-like symptoms which often progress rapidly to fatal complications of hemorrhage, fever, and hypotensive shock. Bowen, E.T.W. et al., *Lancet* 1:571 (1977); Centers for Disease Control, *M.M.W.R.* 15 44:381 (1995); Le Guenno, B. et al., *Lancet* 345:1271 (1995); Peters, C.J. et al., *Fields Virology*, B.N. Fields, D.M. Knipe and P.M. Howley, Eds. (Lippincott-Raven, Philadelphia) p. 1161 (1996). The negative-stranded genome of Ebola virus contains seven structural and regulatory proteins (Sanchez, A. et al., *Virus Res.* 29:215 (1993)), but despite its relative simplicity, the molecular basis for Ebola virus 20 pathogenicity is unknown. Among the viral gene products, the glycoprotein is found in two forms: a secreted form, 50-70 kD (Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996)), synthesized at high levels early in the course of infection, and an alternative transmembrane form, which arises from RNA editing to encode a 120-150 kD glycoprotein that is incorporated into the virion. Sanchez, A. et al., *PNAS (USA)* 25 93:3602 (1996); Volchkov, V.E. et al., *Virology* 214:421 (1995). The first 295 amino acids (aa) of both proteins are identical in the Zaire strain, while sGP contains an additional 69 and GP another 381 COOH- terminal aa residues. Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996). The specific cellular targets of these related gene products and their roles in the pathogenesis of Ebola infection have not been 30 characterized.

SUMMARY OF THE INVENTION

 The present invention provides compositions and methods for targeting gene transfer vectors to certain cell types by pseudotyping with a transmembrane form of viral glycoprotein. In one embodiment, the methods of the invention comprise the

09600766-05403

- 2 -

step of administering to a cell population a gene to be transferred operatively-linked to an appropriate transfer vehicle, wherein the transfer vehicle is associated with a transmembrane form of Ebola glycoprotein. In this embodiment, the gene will be targeted to cell types naturally infected with Ebola such as endothelial cells, monocytes and hepatocytes.

Genetic constructs for selective gene transfer into certain cell types are also provided. The genetic constructs of the present invention comprise a gene to be transferred operatively-linked to an appropriate transfer vehicle or carrier, wherein the transfer vehicle or carrier is associated with a transmembrane form of viral glycoprotein. In one embodiment, the transmembrane form of Ebola glycoprotein is expressed on the surface of a virus-based gene-targeting vector, e.g., lentiviral or retroviral vector. In another embodiment, an expressed or synthesized transmembrane glycoprotein is chemically derivatized to a non-biologic gene targeting vehicle.

Additional objects, advantages, and features of the present invention will become apparent from the following description and appended claims, taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The various advantages of the present invention will become apparent to one skilled in the art by reading the following specification and subjoined claims and by referencing the following drawings.

Figures 1A-1C show the binding of sGP to neutrophils;

Figures 2A-2D show the infection of different cell types by a GP-pseudotyped vector of the present invention;

Figures 3A-3F show the dependence of sGP binding on CD16b and correlation of binding with neutrophil activation;

Figures 4A-4B show the effect of sGP on neutrophil function;

Figures 5A-5C show the infection rate of cells with a GP-pseudotyped retroviral vector of the present invention;

Figure 6 is a schematic of the plasmid pVR 1012-GP(IC) (Ivory Coast strain of GP, see SEQ ID NO: 1);

Figure 7 is a schematic of the plasmid pVR 1012-GP(S) (Sudan strain of GP, see SEQ ID NO: 2);

Figure 8 is a schematic of the plasmid pVR 1012-GP(Z) (Zaire strain of GP, see SEQ ID NO: 3);

- 3 -

Figure 9 is a schematic of the plasmid pVR 1012-sGP(Z) (Zaire strain of sGP, see SEQ ID NO: 4); and

Figure 10 is a summary of the characterization of GP and sGP derivatives for their ability to pseudotype to induce cytotoxicity in producer cells.

5 DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention provides genetic constructs and methods for targeting gene transfer vectors to certain cell types by pseudotyping with a transmembrane form of viral glycoprotein. The methods for selective gene transfer of the present invention comprise the step of administering to a cell population a genetic construct
10 of the present invention so that the gene is transferred and expressed in certain cell types present in the cell population. Administration to the cell population may be *ex vivo* or *in vivo*.

The genetic constructs of the present invention comprise a gene to be transferred operatively-linked to an appropriate transfer vehicle or carrier, wherein the
15 transfer vehicle or carrier is associated with a transmembrane form of viral glycoprotein. In one embodiment, the transmembrane form of Ebola glycoprotein is associated with the vehicle or carrier. The gene to be transferred will thus be targeted to cell types naturally infected with Ebola virus including endothelial cells, hepatocytes, monocytes and related cell types such as dendritic cells. The
20 transmembrane form of Ebola glycoprotein may be chosen from, without limitation, the Ivory Coast strain (SEQ ID NO: 1), Sudan strain (SEQ ID NO: 2), the Zaire strain (SEQ ID NO: 3) and/or the Reston strain. It will be appreciated that in other embodiments of the present invention, other hemorrhagic fever virus glycoproteins, in particular transmembrane glycoproteins, may be employed and will target those cell
25 types naturally infected by the virus. Examples of hemorrhagic viruses include dengue virus, Yellow Fever virus (*flaviviridae*); Lassa, Junin and Machupo (*arenaviridae*); Rift Valley, Congo-Crimean and Hantaan (*bunyaviridae*); and Marburg (*filoviridae*). It will also be appreciated that derivatives of the transmembrane glycoprotein which retain the capability of targeting specific cell types, may also be
30 employed, for example, the transmembrane glycoproteins may be mutated, e.g., toxic regions may be removed to improve producer cell viability (see Figure 10).

The transmembrane glycoprotein may be expressed on the surface of a virus-based gene-targeting vector, e.g., lentiviral, retroviral, replication-deficient retroviral, adenoviral or adeno-associated viral vector. The transmembrane glycoprotein may

09600766-051401

- 4 -

also be expressed or synthesized and chemically derivatized to a non-biologic gene targeting vehicle, e.g., liposome or DNA-protein complex.

The term "operatively-linked" as used herein refers to functional linkage between a nucleic acid expression control sequence (such as a promoter) and a second nucleic acid sequence (*i.e.*, gene), wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence. Expression control sequences are known to those skilled in the art (see, e.g., Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990)). "Associated with" as used herein refers to the transmembrane form of viral glycoprotein being in contact or linkage with the transfer vehicle or carrier in such a way as to direct the transfer vehicle or carrier to certain cell types. The terms "transfer vehicle" and "carrier" refer to any type of structure which is capable of delivering the gene of interest to a target cell.

Many transfer vehicles or carriers are known in the art. For example, various viruses that are capable of infecting cells can be recombinantly manipulated to carry the gene of interest without affecting their infectivity. As used herein, the terms "infect" and "infectivity" refer only to the ability of a virus to transfer genetic material to a target cell. Those terms do not mean that the virus is capable of replication in the target cell. In fact, it is preferable that such viruses are replication defective so that target cells do not suffer the effects of viral replication.

In one embodiment, the virus employed is a replication defective retroviruses. When these replication defective retroviruses are employed, their genomes can be packaged by a helper virus in accordance with well-known techniques. Suitable retroviruses include PLJ, pZip, pWe and pEM, each of which is well known in the art. Suitable helper viruses for packaging genomes include ψ Crip, ψ Cre, ψ 2, ψ Am and adeno-associated viruses.

In another embodiment, lentiviral vectors are employed. Surprisingly, the inventors of the present invention were successful in pseudotyping lentiviral vectors (HIV) with the transmembrane glycoprotein from Ebola. Feline immunodeficiency virus, bovine immunodeficiency virus, simian immunodeficiency virus and EAIV, may also be employed as the carrier in the compositions and methods of the present invention.

Gene delivery systems other than viruses may also be employed. For example, the gene to be transferred may be packaged in a liposome which is chemically derivatized to the transmembrane glycoprotein. To form these liposomes,

- 5 -

one mixes the DNA of an expression vector which expresses the gene to be transferred with lipid, such as *N*-[1-(2,3,-dioleyloxy)propyl]-*N,N,N*-trimethylammonium chloride (DOTMA) in a suitable buffer, such as Hepes buffered saline. This causes the spontaneous formation of lipid-DNA complexes (liposomes). Felgner, P.L. et al.,
5 *PNAS (USA)* 84:7413-7417 (1987).

Another gene delivery system that may be utilized in this invention is DNA-protein complexes. The formation of DNA-protein complexes is described in United States Patent No. 5,166,320, the disclosure of which is herein incorporated by reference.

10 It will be appreciated that any gene may be employed in the compositions and methods of the present invention. For example, and without limitation, in the treatment of cancer, death inducing genes, including genes coding for cytostatic or cytotoxic proteins, e.g., HSV tk, and genes encoding cyclin dependent kinase inhibitors, p21, p27, cytosine deaminase, and fas ligand, may all be employed. In
15 another example, for the treatment of cardiovascular or ischemic vascular disease, genes encoding angiogenic factors such as VEGF basic or acidic FGF's (FGF 1-5) may be employed. In yet another example, in the treatment of vasospasm, the gene encoding NO synthase or heme oxygenase, may be employed. In a further example, monocytes and dendritic cells may be targeted with genes encoding immunogens for
20 cell-targeted immunization.

In one embodiment, the methods of targeting gene transfer vectors to certain cell types involve administering to a cell population *ex vivo*, a construct of the present invention and introducing the transfected cells into a subject. In an alternative embodiment, the methods of the present invention comprise administering to an *in*
25 *vivo* cell population a construct of the present invention. Administration can be by any of the routes normally used for *in vivo* gene therapy such as direct delivery to cells via a gene gun, and other known techniques. The constructs are thus administered in any suitable manner, preferably with pharmaceutically acceptable carriers. The constructs can be administered, for example, by intravenous infusion, orally, topically,
30 intraperitoneally, intravesically or intrathecally. The preferred method of administration will often be intravenous.

To practice an *ex vivo* method of the present invention, a source of cells is obtained. The cells are optionally selected from *in vitro* cells, such as those derived from cell culture and *ex vivo* cells, such as those derived from a subject. The term
35 "subject" is intended to include living organisms, e.g., mammals. Examples of

09600766 "051401
T04T50" 99209960

- 6 -

subjects include humans, primates, dogs, cats, mice, rats, and transgenic species thereof. It will be appreciated that specific cell populations may be obtained by isolation from certain tissues by methods known to those skilled in the art. The cells are maintained under conditions necessary to support growth, for example an appropriate temperature (e.g., 37°C) and atmosphere (e.g., air plus 5% CO₂).

The cells are then transfected with the constructs of the present invention by introducing the constructs to the cell population, under conditions favorable for transfection. According to one embodiment of the present invention, cells are treated with compounds that facilitate uptake of the constructs by the cells. According to another embodiment of the present invention, cells are treated with compounds that stimulate cell division and facilitate uptake of the constructs. It will be appreciated that compounds that facilitate uptake of constructs by cells and compounds that stimulate cell division are known to those skilled in the art.

The constructs of the present invention express the transferred gene in a dose dependent manner. The specific dose to be administered to a patient will be determined by the efficacy of the particular construct and/or delivery system employed, the gene transferred, and the condition of the patient, as well as the body weight or surface area of the patient to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany the administration of a particular construct or effect a particular patient. In determining the effective amount of the construct or transfected cell to be administered, the physician needs to evaluate circulating plasma levels, toxicities, and progression of disease. It will be appreciated that administration can be accomplished via single or divided doses.

There is a wide variety of suitable formulations for pharmaceutical compositions containing the constructs of the present invention. Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the construct dissolved in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the construct, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. The construct, alone or in combination with other suitable components, may also be made into aerosol formulations to be administered via inhalation, e.g., to the bronchial passageways. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

- 7 -

Suitable formulations for rectal administration include, for example, suppositories, which consist of the construct with a suppository base. Suitable suppository bases include natural or synthetic triglycerides or paraffin hydrocarbons. In addition, it is also possible to use gelatin rectal capsules which consist of a combination of the construct with a base, including, for example, liquid triglycerides, polyethylene glycols, and paraffin hydrocarbons.

Formulations suitable for parenteral administration, such as, for example, by intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The formulations can be presented in unit-dose or multi-dose sealed containers, such as ampules or vials. Extemporaneous injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. Cells transfected by the constructs as described above in the context of *ex vivo* therapy can also be administered as described above.

This invention also provides compositions and kits comprising the constructs of the present invention. For example, the composition can comprise the constructs of the present invention in a pharmaceutically acceptable carrier as described above. Kits comprising such compositions and instructions for use are also within the scope of this invention.

In order to more fully demonstrate the advantages arising from the present invention, the following examples are set forth. It is to be understood that the following is by way of example only and is not intended as a limitation on the scope of the invention.

SPECIFIC EXAMPLE 1

I. Methods

Recombinant retroviruses were produced by transient transfection of 293T cells: 2×10^6 cells were plated 24 hours before transfection in 60 mm dishes. Transfection was performed by calcium-phosphate precipitation using 3 μ g of a retroviral vector (Kinsella, T.M. et al., *Hum. Gene Ther.* 7:1405 (1996)) encoding luciferase linked to an internal ribosome entry site and a green fluorescent protein derivative (GFP; pEGFP, Clontech), pLZRs-Luc-Gfp, 5 μ g of an expression vector

- 8 -

encoding gag and pol, pNGVL-MLVgag-pol, and 1 μ g of the envelope encoding plasmid: pNGVL-4070A (ampho) env, pCMV-Eco env or p1012-Ebola GP, respectively. Supernatants corresponding to 24-48 hours post-transfection were harvested, cleared by low-speed centrifugation and either used immediately for infection or frozen at -80°C. Infections were performed in 6-well plates ($1.5-2.5 \times 10^5$ adherent cells) or 12-well plates (5×10^5 non-adherent) using different dilutions of the supernatants by incubating the cells overnight with 1 ml and 300 μ l, respectively of the diluted supernatants. Polybrene was used at a concentration of 5 μ g/ml for all the cell lines except for D17 in which the concentration was 100 μ g/ml. After overnight infection, fresh medium was added and the cells were incubated for an additional 24 hours. After infection, the cells were lysed in 25 mM Tris-phosphate pH 8, 2 mM DTT, 2 mM 1,2-diaminocyclohexene-N,N',N'-tetraacetic acid, 10% glycerol, 1% TritonX-100, and assayed for luciferase activity using Luciferase Assay Reagent (Promega, Madison, WI) in a 1251 BioOrbit Luminometer. The same number of cells (range $5-10 \times 10^4$) was analyzed for every specific cell line.

Binding of sGP to neutrophils and inverse correlation of binding with activation: Figures 1A-1A2. PBMC from normal volunteers were incubated with control or sGP supernatants derived from transfected 293 cells, and immunostaining was performed using a rabbit antibody to sGP as previously described. Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996); Xu, L. et al., *Nat. Med.* (1997) in press. Secondary staining was performed with a fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit IgG antibody (Sigma, F9887). All incubations were performed at 4°C for 30 minutes with .4 μ g of the relevant antibodies per 10^6 cells in a 50 μ l volume.

Figures 1B-1B1. Double immunostaining with antibodies to sGP and the neutrophil-specific marker, CD15. Cells were incubated with a FITC conjugated mouse anti-human CD15 antibody (Caltag, cat# MHCD1501), followed by secondary staining with a PE-conjugated anti-rabbit IgG antibody (Sigma) to detect sGP binding. Cells were washed with PBS, fixed in 1% formaldehyde, and analyzed by FACS.

Figure 1C. Specific absorption of sGP by neutrophils. Control or sGP supernatants derived from relevant transfected 293 cells (Xu, L. et al., *Nat. Med.* (1997) in press) were incubated at 1:500 dilution with 10^6 mononuclear or granulocytic cells. Cells were removed and the resulting supernatants analyzed by an 8% SDS PAGE gel. Western blot analysis was performed as previously described (Xu, L. et al., *Nat. Med.* (1997) in press) using an anti-GP rabbit antisera and a secondary

antibody, horseradish peroxidase conjugated donkey anti-rabbit IgG at a dilution of 1:5,000 (Amersham, NA934). Primary antibody was incubated for 30 minutes at room temperature, as was the secondary antibody. The immunocomplexes were detected by chemiluminescence using Supersignal® chemiluminescent substrate reagents (Pierce) according to the manufacturer's instructions. Arrow indicates sGP reactive band.

Infection of different cell types by GP-pseudotyped retroviral vector and preferential binding to endothelial cells: Figure 2A. Infection of different indicator cell lines with the Ebola-GP pseudotyped retrovirus expressing luciferase.

Amphotropic and ecotropic retroviral vectors were used as controls. Viruses were diluted to different multiplicities of infection (MOI) to provide for equal luciferase activity on Hela cervical epithelial cells, permissive for amphotropic retrovirus, D17 dog osteosarcoma cells (Embretson, J.E. et al., *J. Virol.* 61:3454 (1987)), which are permissive for amphotropic, xenotropic, and ecotropic retroviruses, and BW5147 T leukemia cells permissive for amphotropic and ecotropic virus. In these groups, GP virus titer was $1-4 \times 10^5$ /ml and amphotropic virus was $\sim 2 \times 10^4$ /ml (MOI's ≈ 1.0 and 0.1 , respectively), and the ecotropic virus titer was $\sim 10^6$ /ml (MOI ≈ 10). Titers were determined by endpoint dilution of reporter activity of the amphotropic virus in D17 cells, and was normalized to reverse transcriptase activity for the GP virus.

Figure 2B. Analysis of different normal or transformed cell lines by infection with amphotropic or GP retroviral vectors at the same titer (10^4 /ml, MOI ≈ 0.2). Forty-eight hours after infection, an equivalent of 5×10^4 cells was assayed for luciferase activity after exposure to equal titers of viral stocks. Luminescence is expressed as the fold-increase over non-infected control cells.

Figures 2C-2C3. The binding of sGP (left) or GP-pseudotyped retrovirus (right) to neutrophils (upper panel) or microvascular endothelium (lower panel) was determined by FACS. sGP binding was performed as in Fig. 1A, and retrovirus incubation was performed at 37°C for 2 hours in the presence of polybrene ($8 \mu\text{g}/\text{ml}$).

Figure 2D. Infection of D17 cells by GP-pseudotyped virus in the absence (lane 1, none) or presence of control (lane 2) or sGP supernatant (lane 3) from transfected 293 cells. Gene transfer was measured by the luciferase assay as described below. Luminescence refers to relative light units in the luciferase assay.

Dependence of sGP binding on CD16b and correlation of binding with neutrophil activation: Figures 3A-3D. Neutrophils were incubated for 30 minutes at 4°C with a mouse antibody to CD16b (upper panel; clone 3G8 from Immunotech,

- 10 -

cat# 1M0813) or CD62L (middle panel, R&D Systems), compared to the indicated control antibody [purified mouse IgG (Vector Laboratories), I-2000], followed by supernatants from control or sGP-transfected 293 cells, primary rabbit antibody to sGP, and a FITC-conjugated secondary antibody to rabbit IgG (Fig. 1, legend). Cells were washed with PBS, fixed in 1% formaldehyde, and analyzed by FACS. For blocking, 10^6 cells were incubated with 0.5 – 1 μ g of the relevant antibodies for 30 minutes in a 50 μ l volume.

Figures 3E-3F. Immunostaining with sGP was performed on isolated neutrophils which were maintained in media (none) or incubated with PMA (10 ng/ml) at 37°C for 30 minutes (PMA).

Effect of sGP on neutrophil function: *Figures 4A-4B.* Exposure of neutrophils to sGP inhibits down modulation of L-selectin. Isolated neutrophils were incubated with the indicated control or sGP containing supernatants (Xu, L. et al., *Nat. Med.* (1997) in press) and defined media (AIM V, GIBCO) for 4 hours at 37°C. Expression of L-selectin was determined using an anti-CD62L antibody (R&D Systems), followed by the secondary staining using a FITC-conjugated anti-mouse IgG (Sigma, F2883) as described in Fig. 1, legend. Cells were washed with PBS, fixed with 1% formaldehyde and analyzed by FACS for relative levels of fluorescence intensity as a function of cell number. An isotype control was used to quantitate background levels of immunostaining (neg.). Results are representative of three independent experiments.

II. Results

To determine the specificity of Ebola virus glycoproteins, expression vectors encoding either sGP, GP, or a plasmid control (Xu, L. et al., *Nat. Med.* (1997) in press) were transfected into 293 cells, and cell culture supernatants were used as a source of relevant recombinant glycoproteins. Binding of sGP was determined by immunofluorescence analysis after incubation of relevant supernatants with normal or transformed human cell lines. No binding was detected to several hematopoietic lineages, including lymphocytes or monocytes (Fig. 1A), or transformed Jurkat or CEM T leukemias, the HL60 myelomonocytic or U937 promonocytic leukemia cells. In contrast, sGP was able to bind to granulocytic cells, as evidenced by FACS analysis of this subset of peripheral blood mononuclear cells (PBMC) discriminated by cell size and granularity (Fig. 1A). This cell specificity was confirmed by using double-staining with a granulocyte-specific cell surface marker, CD15 (Fig. 1B). Absorption of sGP

- 11 -

by purified neutrophils in the absence of antibodies also resulted in depletion of sGP, indicating that binding to the neutrophil occurred in the absence of antibody (Fig. 1C).

A potential structural similarity between Ebola GP and avian sarcoma virus envelope protein has been previously proposed (Gallaher, W.R., *Cell* 85:477 (1996)),

5 raising the possibility that this protein could be incorporated into retroviral particles.

To determine the binding specificity of the transmembrane glycoprotein, pseudotyping of a Moloney leukemia virus was therefore attempted. Infectivity of different cell types by this pseudotyped vector was determined with a luciferase reporter gene. de Wet,

J.R. et al., *Mol. Cell. Biol.* 7:725 (1987). This analysis revealed infection of cells

10 different from those which interacted with sGP (Fig. 2A,B). For example, though it

could infect other cell types, transduction by the GP retroviral vector readily occurred

in endothelial cells, either from the microvasculature (MVEC) or umbilical veins

(HUVEC) (Fig. 2B), which did not bind sGP (Fig. 2C, left). When the specificity of

GP-retrovirus was compared to murine retroviruses pseudotyped with amphotropic or

15 ecotropic envelope gp70 proteins, the range of susceptible target cells differed (Fig.

2B), suggesting that the virus receptor(s) for Ebola GP differ from those previously

described for gp70. Minimal binding of GP-virus was observed on neutrophils, despite

the ability of these cells to bind sGP (Fig. 2C, upper panel) and the fact that

immunoreactive protein was detected on the virus. Conversely, GP-virus binding to

20 endothelial cells was readily detected, though these cells did not bind sGP (Fig. 2C,

lower panel). When sGP was analyzed for its effect on GP retroviral gene transfer,

infection was not inhibited by sGP (Fig. 2D), further confirming the divergent

specificities of the two forms of the viral glycoprotein. Recent studies have revealed

that the biochemical forms of these proteins differ, with sGP present in solution

25 primarily as a homodimer and GP as a trimer, suggesting that differences in multimer

composition may contribute to these alternative specificities.

Potential cell surface receptors for sGP were analyzed with antibodies to several neutrophil cell surface antigens to interfere with sGP binding, including CD15,

L-selectin (CD62L), CD16b, and several common leukocyte antigens. Only the

30 neutrophil-specific form of the low affinity $F_c \gamma$ receptor III, CD16b, inhibited sGP

binding specifically. Antibodies to CD62L, for example, did not inhibit sGP binding

(Fig. 3). Binding to neutrophils correlated with their activation state and CD16b

expression since no binding was observed in cells stimulated with phorbol 12-

myristate 13-acetate (PMA) for 30 minutes, at which time CD16b expression was

35 markedly decreased on these cells (Fig. 3, lower panel). Overexpression of this form

T04750-99200960

- 12 -

of CD16 on a heterologous cell type, 3T3 fibroblasts, did not confer sGP binding to these cells by FACS analysis, suggesting that CD16b is necessary but not sufficient for stable binding.

Binding of sGP did not inhibit neutrophil activation in response to potent pleiotropic activators (PMA, IL-8, or f-Met-Leu-Phe), as measured by down modulation of L-selectin expression using FACS analysis. In a defined serum-free medium, partial activation of neutrophils was observed, with a decrease in L-selectin expression at 4 hours (Fig. 4). Under these conditions, incubation of neutrophils with sGP supernatant prevented this decrease in L-selectin expression (Fig. 4). Because L-selectin was not required for sGP binding (Fig. 3), this effect was apparently indirect, through a mechanism not yet defined, possibly involving CD16b or carbohydrate interactions of the highly glycosylated sGP protein.

The expression of alternative Ebola virus glycoproteins in clinical infection has long been recognized, but their functional roles and cell specificity have not been defined. Early after infection, high levels of the secreted protein are found in the serum and precede fulminant replication and dissemination of virus systemically, at which time synthesis of transmembrane GP is markedly increased. Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996). The inventors have now found that the binding specificities of these two molecules differ. It had been proposed that sGP may serve as a decoy to prevent recognition of GP, possibly to temporarily inhibit virus binding to target cells. The studies set forth herein suggest that this hypothesis is unlikely to be correct. The binding specificities of these proteins differ, and despite the fact that they are derived from the same viral gene, it has been surprisingly found that alternative forms of the glycoprotein have been selected for different functions.

Although these proteins share identical amino terminal sequences, their carboxyl terminal regions differ. Sanchez, A. et al., *Virus Res.* 29:215 (1993). These sequences are likely responsible for the differences in binding specificity, either through direct interactions in these domains or by their effect on multimerization. The secreted glycoprotein binds to neutrophils to prevent early events in activation, possibly serving to diminish any inflammatory responses which might provide innate immunity to the virus, facilitating productive viral replication. The subsequent increase in GP synthesis gives rise to virus which in turn could infect other cells. Filoviruses have been shown previously to infect and replicate in different cell types and appear to grow readily in endothelial cells *in vivo*. Peters, C.J. et al., *Fields Virology*, B.N. Fields, D.M. Knipe and P.M. Howley, Eds. (Lippincott-Raven, Philadelphia) (1996);

Schnittler, H.J. et al., *J. Clin. Invest.* 91:1301 (1993). The findings set forth herein suggest that its tropism for this cell type is probably determined by the specificity of GP. In Ebola infection, preferential binding and infection of microvascular endothelial cells may lead ultimately to a loss of capillary integrity that results in the severe hemorrhage observed in the terminal stages of this disease. The differential binding of these two gene products from the same viral structural gene generated by RNA editing suggests that they have evolved functionally to differentially affect immunity and infectivity. The ability to facilitate viral replication and target the virus to endothelial cells by alternative products of the same viral gene represents an efficient genetic mechanism which can account for different pathologic features of this disease. Inhibition of sGP binding to neutrophils and GP to endothelium is likely to ameliorate the effects of acute Ebola virus infection.

SPECIFIC EXAMPLE 2

I. Methods

Production of pseudotyped MuLV retroviruses expressing green fluorescent protein (GFP): 50% - 70% confluent 293 T cells in 60mm tissue culture dishes were transfected using the calcium phosphate method and the following plasmids: 0.3 μ g 1012 GP(Z) (see Figure 8) or 1012 sGp-Gp(Z) (see Figure 9), 3 μ g LZR-gfp, 2 μ g pNGVL-gag-pol. After overnight transfection, fresh media was added to cells. Twenty hours later, the supernatants were harvested and filtered through a .45 μ m filter.

Infection of HUVEC cells using the pseudotyped retroviruses: The day before infection, 30% - 50% confluent HUVEC cells were prepared in 6-well plates. 1 ml of pseudotyped retroviral supernatant was added to one well of the 6-well plates with 15 μ g/ml of polybrene. Sixteen hours later, the viruses were removed and normal media was added. After 24 hours, the cells were lifted and GFP expression measured using FACS analysis.

Construction of 1012 sGP-GP(Z): 1012 sGP(Z) cells were digested with PstI and treated with Klenow, then digested with XbaI. 1012 GP(Z) cells were digested by EcoRI and treated with Klenow, then digested with KpnI. PstI/Klenow/XbaI treated sGP fragment and EcoRI/Klenow/KpnI treated GP fragment were then cloned into XbaI/KpnI treated pVR-1012 plasmid.

GP and sGP derivatives: The receptor recognition domain, mucin-like domain and/or TM domain of GP and sGP were mutated. The mutated GP and sGP was then tested for its ability to pseudotype and for cytotoxicity in producer cells.

II. Results

To determine the efficacy of targeting endothelium with the gene transfer vectors pseudotyped with GP of the present invention, HUVEC cells were infected with GP(Z) pseudotyped MuLV retrovirus (LZR-gfp) and sGP-GP(Z) pseudotyped
5 MuLV retrovirus (LZR-gfp). Figures 5A-5C show the infection rate (GFP expression) measured using FACS analysis. As shown in Figure 5B, the GP(Z) pseudotyped MuLV retrovirus (LZR-gfp) was effective in targeting and expressing GFP in endothelium.

To determine whether mutating GP would effect its ability to pseudotype and/or
10 decrease toxicity in producer cells, the receptor recognition domain, mucin-like domain and/or TM domain were mutated. Figure 10 shows the results. The optimal envelope is able to pseudotype but shows minimal toxicity.

The foregoing discussion discloses and describes merely exemplary embodiments of the present invention. One skilled in the art will readily recognize
15 from such discussion, and from the accompanying drawings and claims, that various changes, modifications and variations can be made therein without departing from the spirit and scope of the invention as defined in the following claims.

All patents and other references cited herein are incorporated by reference as if fully set forth.

09600766-051401

WE CLAIM:

1. A genetic construct comprising a gene operatively-linked to a carrier, wherein the carrier is associated with a transmembrane form of viral glycoprotein or derivative thereof.

5 2. The genetic construct of Claim 1, wherein the transmembrane form of viral glycoprotein or derivative thereof is expressed on the surface of the carrier.

3. The genetic construct of Claim 1, wherein the transmembrane form of viral glycoprotein or derivative thereof is from Ebola.

4. The genetic construct of Claim 1, wherein the carrier is a viral vector.

10 5. The genetic construct of Claim 1, wherein the carrier is a non-biologic gene targeting vehicle.

6. The genetic construct of Claim 4, wherein the viral vector is a retroviral vector.

15 7. The genetic construct of Claim 4, wherein the viral vector is a lentiviral vector.

8. The genetic construct of Claim 5, wherein the non-biologic gene targeting vehicle is a liposome.

9. The genetic construct of Claim 5, wherein the non-biologic gene targeting vehicle is a DNA-protein complex.

20 10. A method of targeting a gene to a cell comprising the step of administering to a cell population a genetic construct comprising the gene operatively-linked to a carrier, wherein the carrier is associated with a transmembrane form of viral glycoprotein or derivatives thereof.

25 11. The method of Claim 10, wherein the transmembrane form of viral glycoprotein or derivative thereof is expressed on the surface of the carrier.

09600766-051401

- 10 20. The method of Claim 14, further comprising the step of introducing the cell population to a subject.

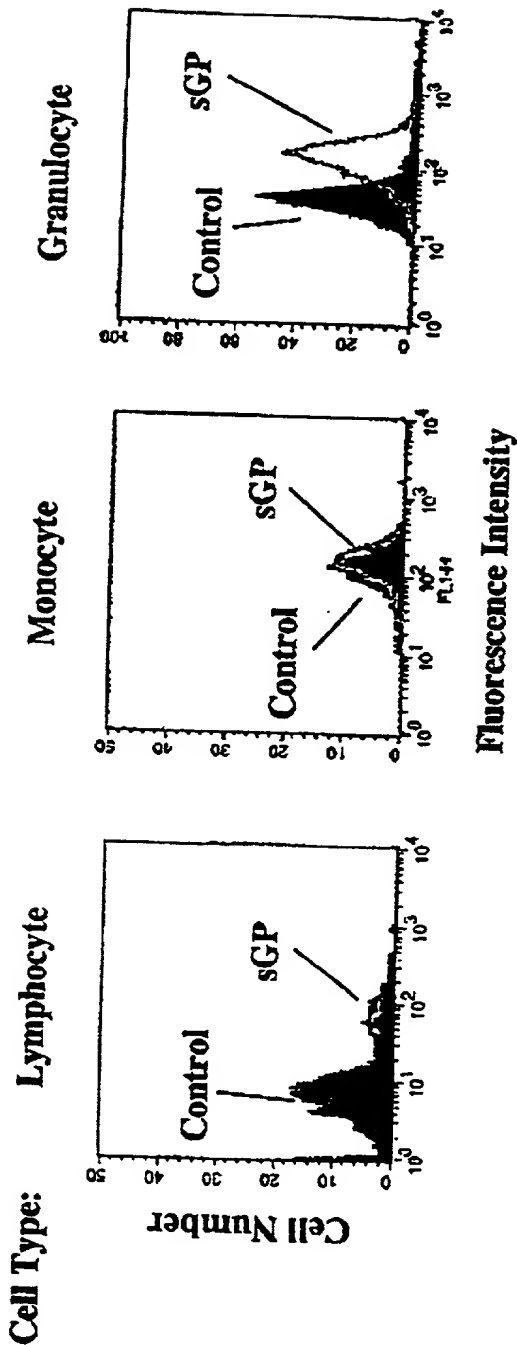


Figure 1A

Figure 1A1

Figure 1A2

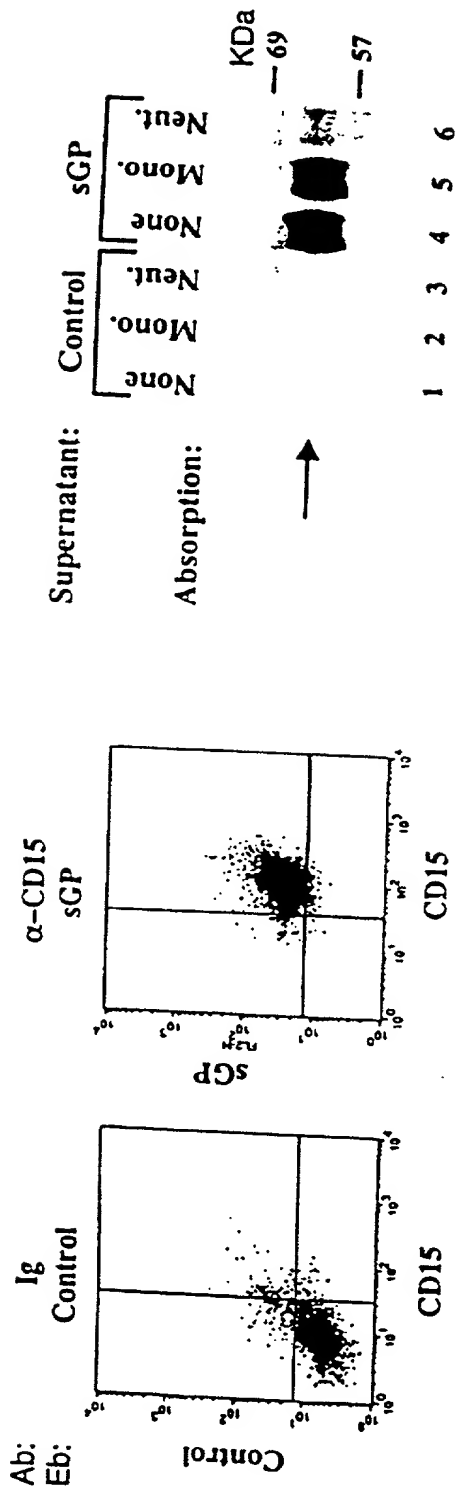


Figure 1B

Figure 1B1

Figure 1C

2/11

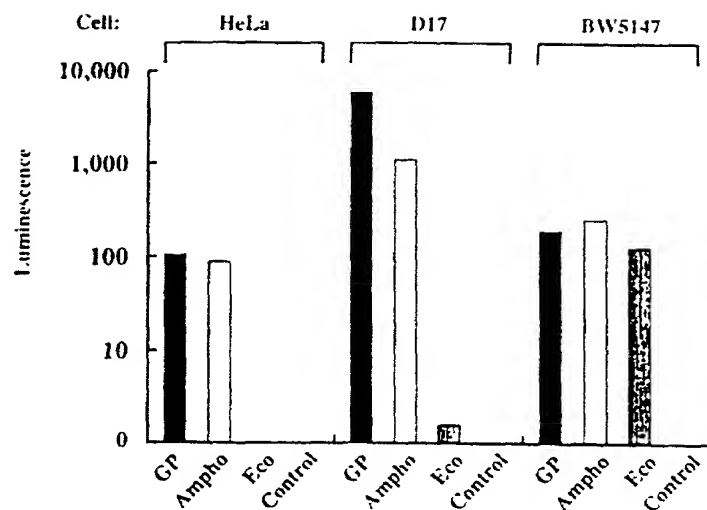


Figure 2A

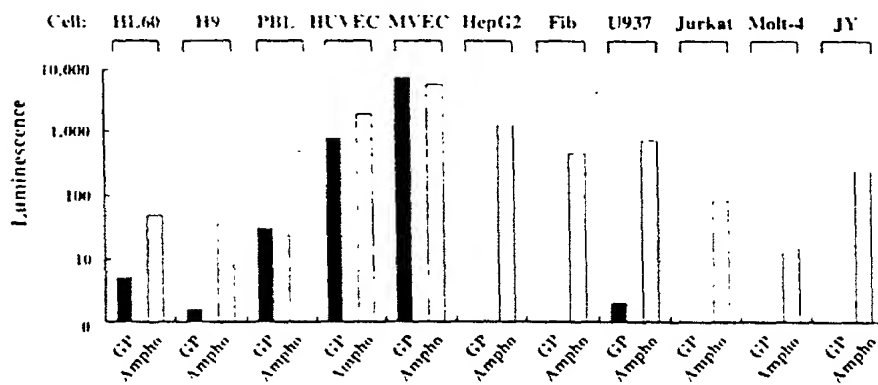


Figure 2B

104150" 33200960

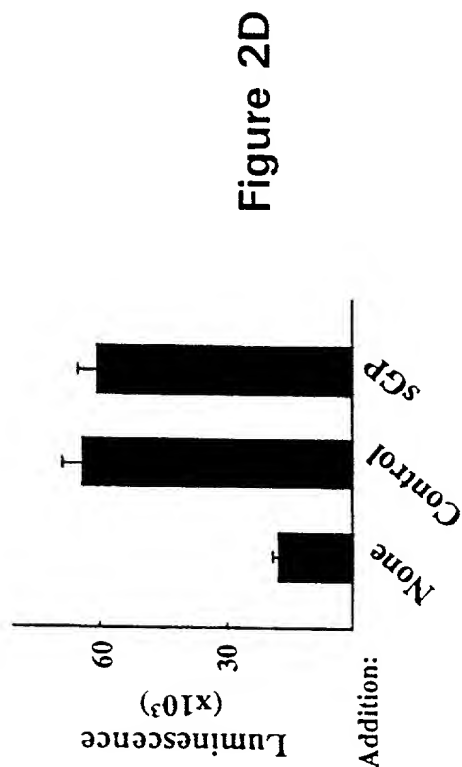
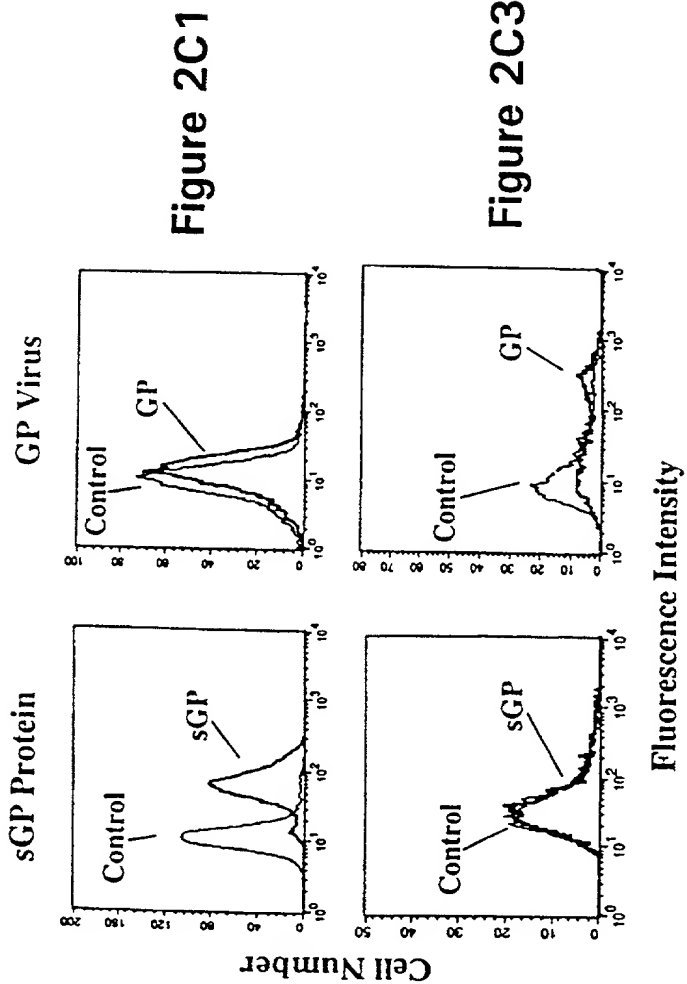


Figure 3A

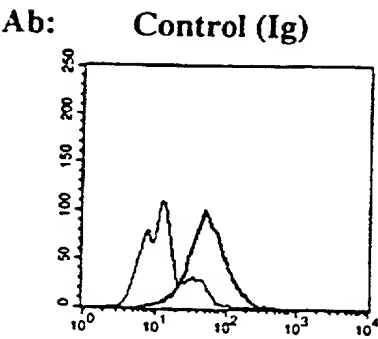


Figure 3B

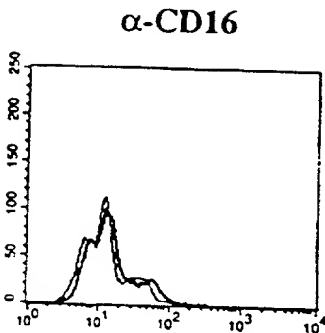


Figure 3C

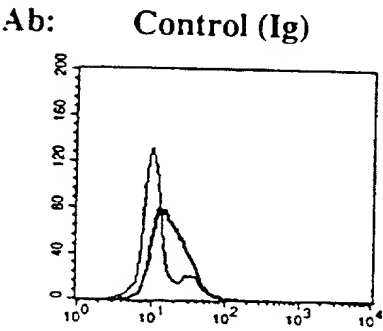


Figure 3D

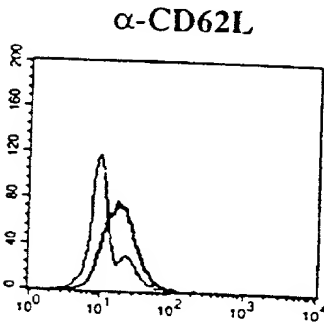


Figure 3E

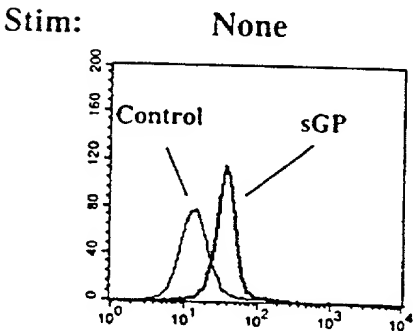
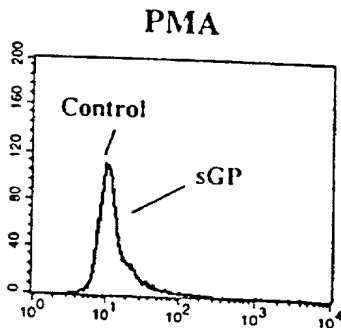


Figure 3F



Fluorescence Intensity

Cell Number

T04T50" 99/00960

5/11

Figure 4A

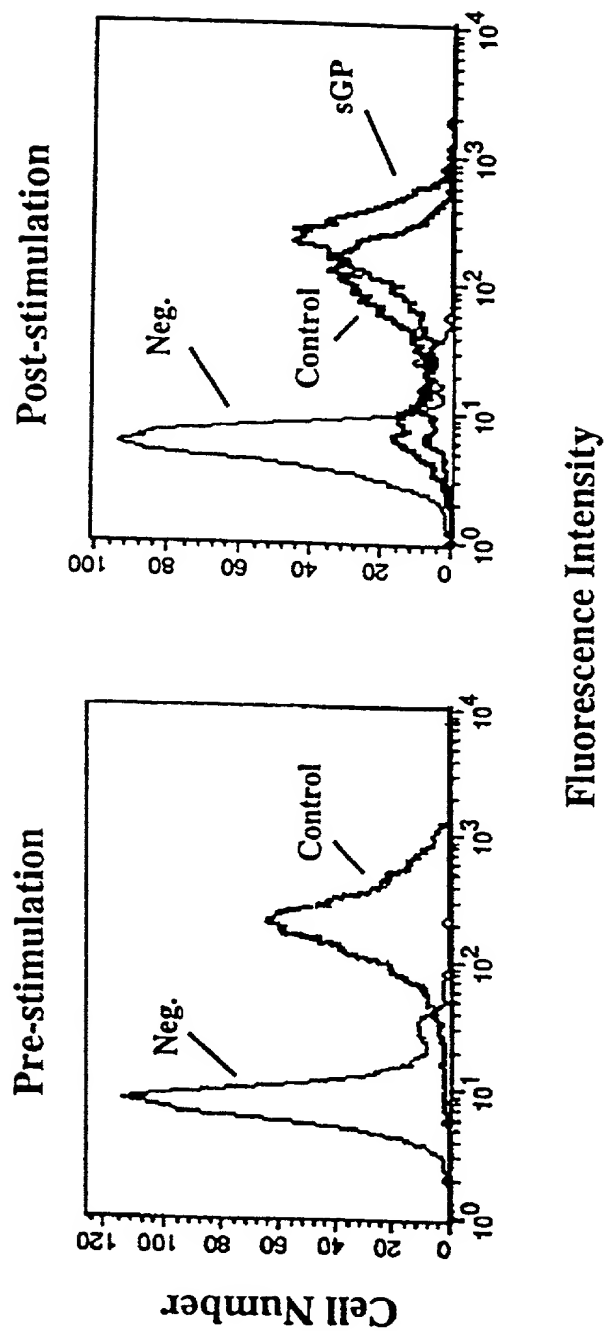


Figure 4B

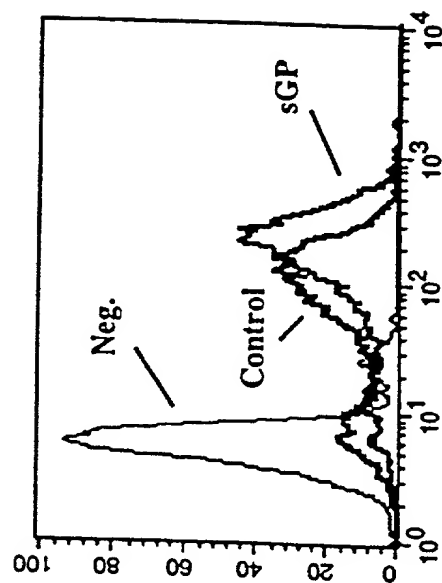


Figure 5A

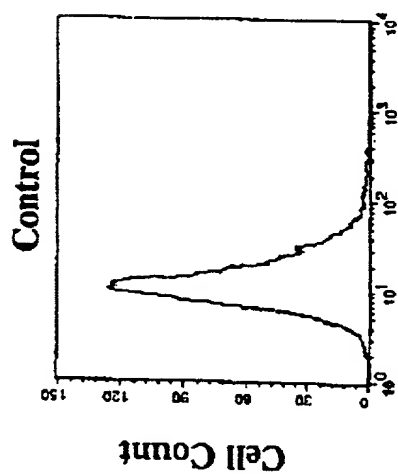


Figure 5B

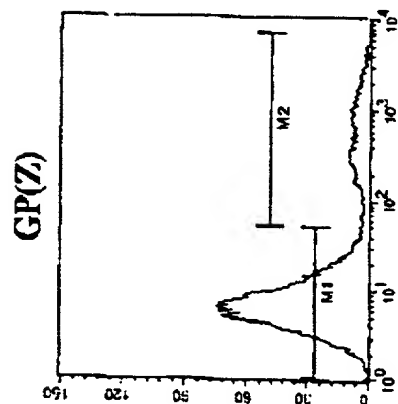
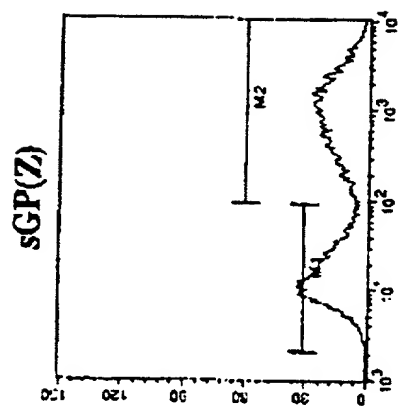


Figure 5C



7/11

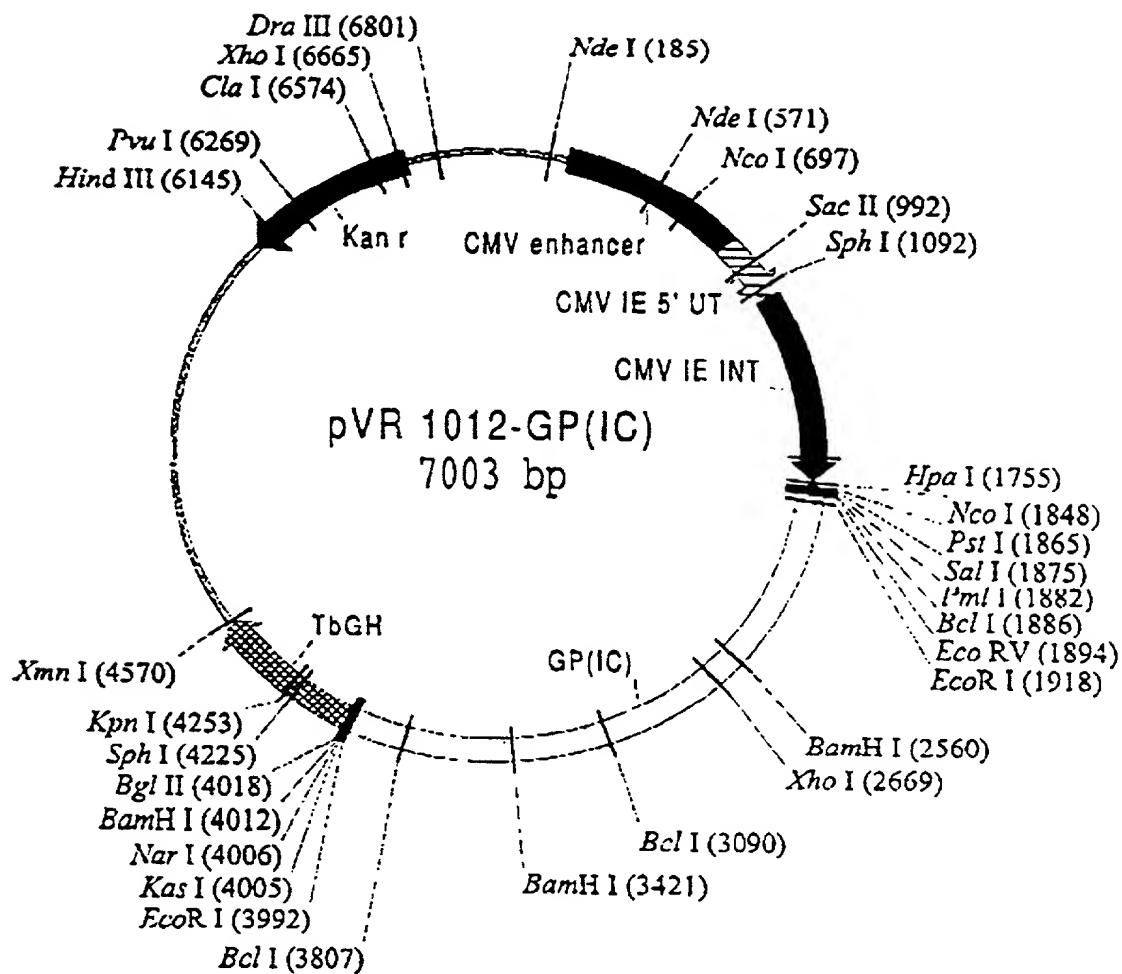


Figure 6

06600766-051401

8/11

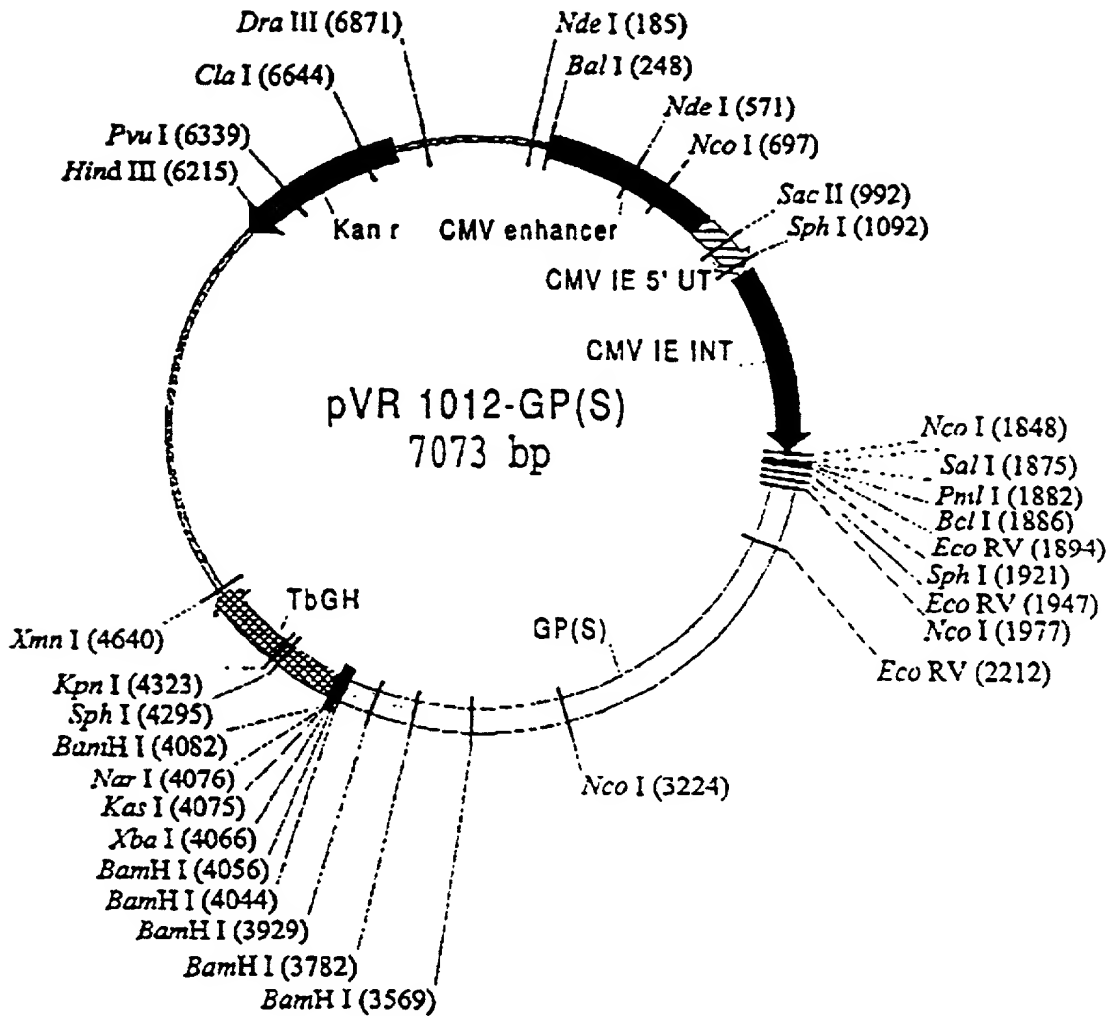


Figure 7

9/11

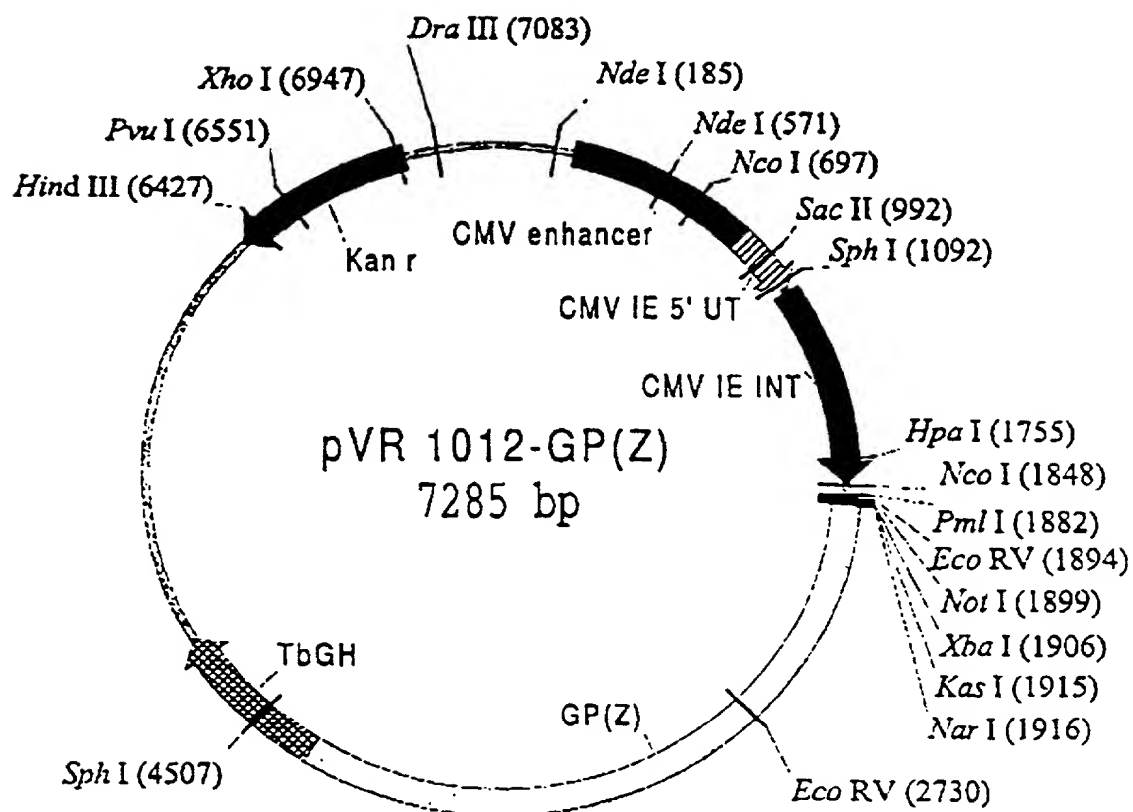


Figure 8

10/11

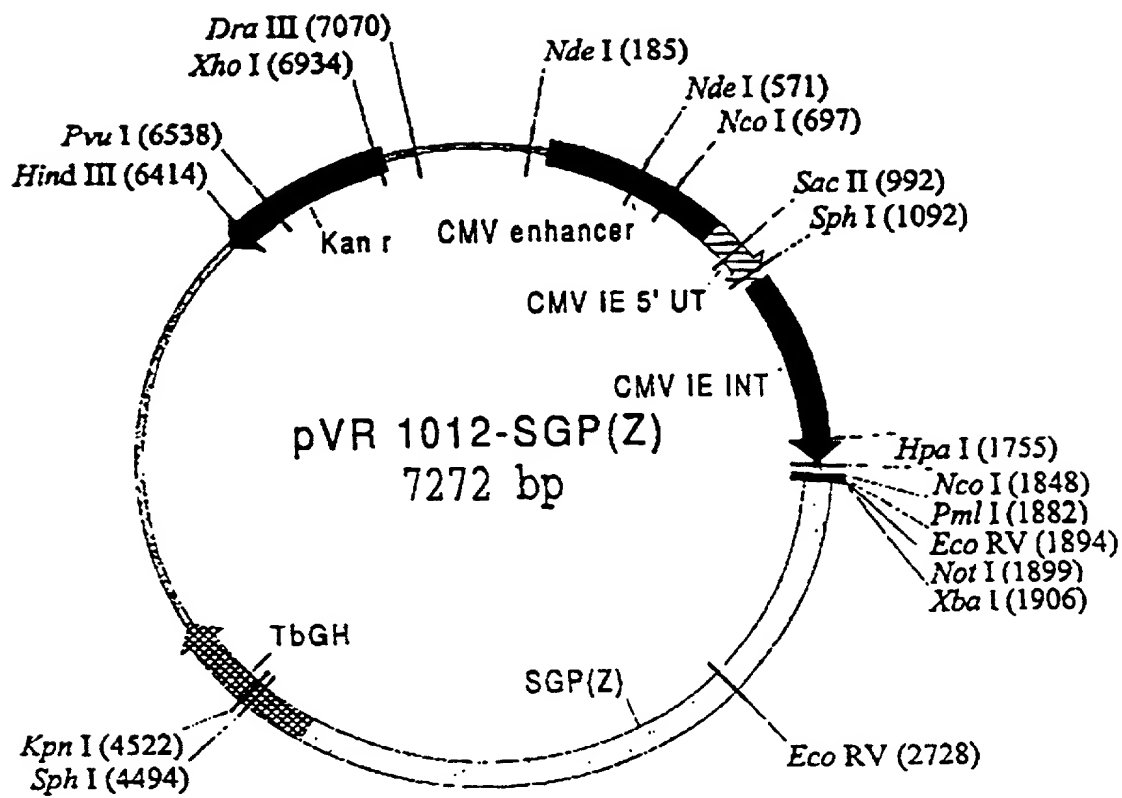


Figure 9

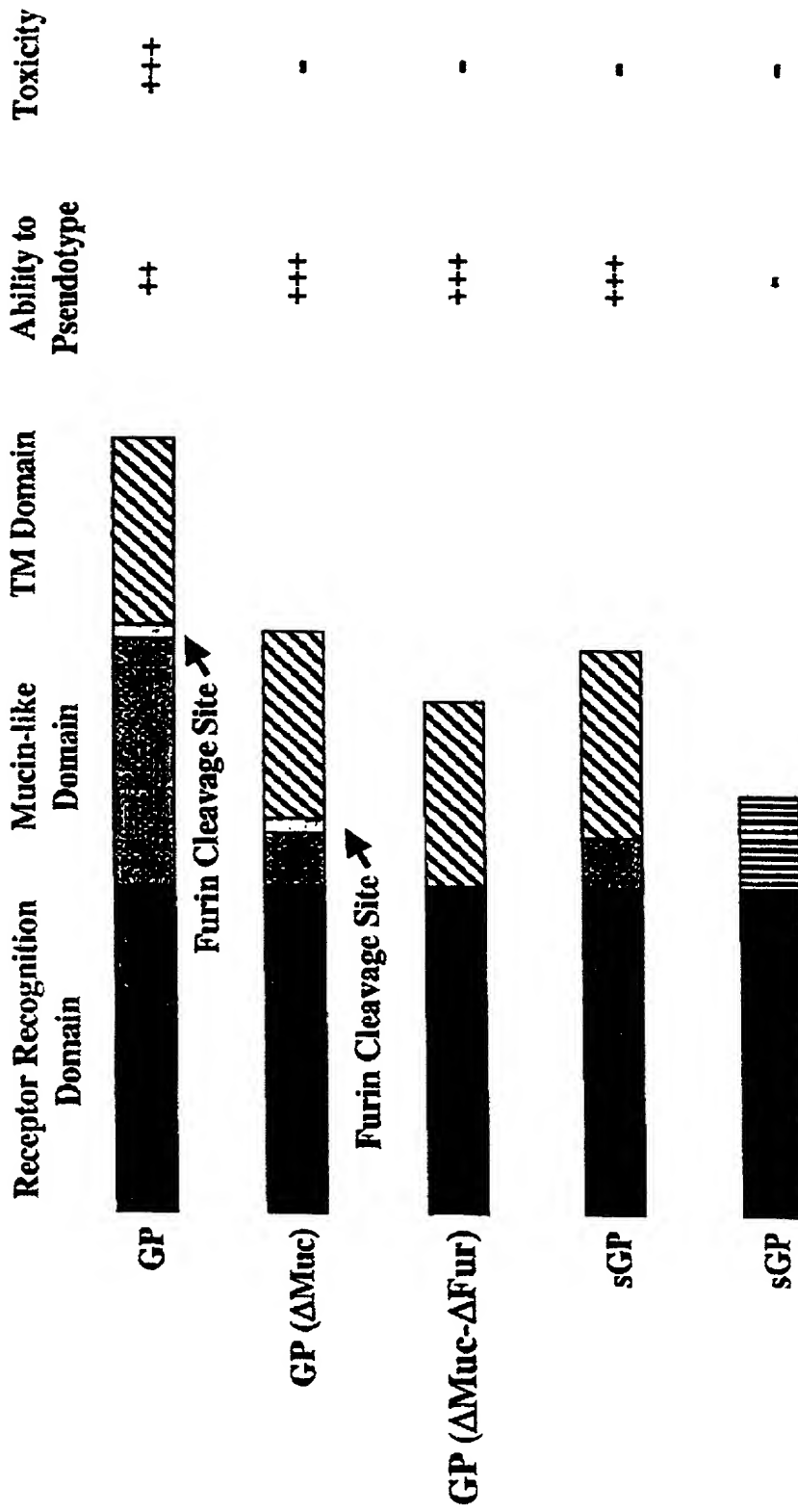


Figure 10

SEQUENCE LISTING ID NO: 1

pVR 1012-GP (IC)

General Description

DNA pVR 1012-GP (IC)
 Local object
 Created: 09/14/98 04:17PM
 Last Modification Date: ? (no data)
 length: 7003 bp
 storage type: Basic
 form: Circular

Comments

Restriction Map

BglII: 1 site AGATCT
 TCTAGA
 ClaI: 1 site ATTCAT
 TAGCTA
 DraIII: 1 site CACTTNGTG
 GTGNNNCAC
 EcoRV: 1 site GATATC
 CTATAG
 HindIII: 1 site AAGCTT
 TTCGAA
 HpaI: 1 site GTTAAC
 CAATTG
 KsaI: 1 site GCGCGC
 CCGCGG
 KpnI: 1 site GGTACC
 CCTTGG
 NarI: 1 site GCGCGC
 CCGCGG
 PmlI: 1 site CACGTG
 GTGCAC
 PstI: 1 site CTGCAG
 GAGCTC
 PvuI: 1 site CGATCG
 GCTAGC
 SacII: 1 site CCGCGG
 GCGGCC
 SalI: 1 site GTCGAC
 CAGCTG
 XmnI: 1 site GAANNNTTC
 CTTNNNAAG
 EcoRI: 2 sites GAATTC
 CTTAAG
 NcoI: 2 sites CCATGG
 GGTACC
 NdeI: 2 sites CATATG
 GTATAC
 SphI: 2 sites GCATGC
 CGTACG
 XhoI: 2 sites CTCGAG
 GAGCTC
 BamHI: 3 sites GGATCC
 CCTAGG

T04T50-99200960

BclI: 3 sites TCATCA
 ACTACT

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4020 End: 4572

Kan r

Start: 6068 End: 6690 (Complementary)

Misc_feature (2 signals)

CMV enhancer

Start: 248 End: 885

GP(IC)

Start: 1870 End: 4019

Annotations

1 TCCGCGGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT CCAGCTCCCG
AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

51 GAGACGGTCA CAGCTTGTCT GTAAGCGAT CCGCGGAGCA GACAAGCCCG
CTCTGCCAGT GTCGAACAGA CATTGGCTA CGGCCCTCGT CTGTTGGGGC

101 TCAGGCGCGG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTAACTATG
AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCGGACC GAATTGATAC

NdeI

151 CGGCATCAGA GCAGATTCTA CTGAGAGTCC ACCATATGCG GTGTCAAATA
GCCGTAGTCT CGTCTAACAT GACTCTCAGC TGGTATACGC CACACTTTAT

201 CCGCACAGAT CGGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
GGCGTGCTA CGCATTCTCT TTTTATGGCG TAGTCTAACG GATAACCGGT

251 TTGCATACGT TGTATCCATA TCATAATATG TACATTATA TTGGCTCATG
AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

301 TCCACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATACT
AGTTTGTAAAT GCGGGTACAA CTGTAATAA TAACTGATCA ATAATTATCA

351 AACCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CARGGCGCAA

401 ACATAACTTA CGGTAATGCG CCGGCTGGC TGACCGCCCA ACGACCCCGG
GTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGCC

451 CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
GGGTAATGCG AGTTATTACT GCATACAAGG GTATCATTCG GTTTATCCCT

501 CTTTCCATTG ACGTCAATCG GTGGASTATT TACGGTAAAC TGCCCACTTG
GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTC ACGGGTGAAC

NdeI

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
CGTCATGTAG TTCACATAGT ATACGGTTCA TCGGGGGGAT AACTCCACTT

601 TGACGGTAAA TGCCCCGCGT GGCATTATGC CCACTACATG ACCTTATGGG
ACTGCCATT TACGGGCGGA CCGTAATACG GGTCAATGAC TCGAATACCC

NcoI

651 ACTTTCTTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
TGAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCC ATAATGGTAC

NcoI

701 GTGATGCGGT TTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACAGAC

751 ACGGCGATTT CCAAGTCTCC ACCCATTTGA CGTCAATGGG AGTTTGTTTT
TGCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAAA

801 GGCACCAAAA TCAACGGGAC TTTCAAAAAT GTCGTACAA CTCCGCCCCA
CCGTGTTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTTGT GAGGCGGGGT

851 TTGACGCAAA TGGGCGGTAG GCGGTACCG TGGGAGGTCT ATATAAGCAG
AACTGCGTTT ACCCGCCATC CCGACATGCC ACCCTCCAGA TATATTGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGT
TCGAACAAAT CACTTGCGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CCGGACCGAT CCAGCCTCCG CGCCCGGGA
AACTGGAGGT ATCTTCTGTG GCCCTGCGTA GGTGCGAGGC CGCGGCCCTT

1001 CCGTGCAATTG GAACCGCGAT TCCCCGTCCT AAGAGTGACC TAAGTACCGC
GCCACGTAAC CTTGCGCCTA AGGGGCGCGG TTCTCACTCC ATTCAATGCG

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTAAGCAT GCTATACTGT
GATACTGAG ATATCCGTGT GGGGAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG GCGCCTATAC ACCCCCGCTT CTTATGCTA TAGGTGATGG
AAAACCGAAT CCCCGATATC TCGGGCGGAA GGAATACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGCTAATAAC TGGTGAGGGG

1201 TATTGGTGAC CAGACTTTC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTG CATTGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CAACATCTTC TATTGGCTAT ATGCCAATAC TCTGTCTTC AGAGACTGAC
GTTCATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTACA CGATGGGGTC CCATTTATTA TTTACAAAT
TGCCTGAGAC ATAAAAATGT CCTACCCAG GGTAAATAAT AATGTTTAA

1351 CACATATACA ACAACGCGT CCCCCGTGCC CGCAGTTTCT ATTAACATA
GTCTATATGT TGTTCGCGCA GGGGCGACGG CGGTCAAAA TAATTTGTAT

1401 GCGTGGGATC TCCACCGGAA TCTCGGTAC GTGTCCGGA CATGGGCTCT
CGCACCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA

1451 TCTCCGGTAG CGCGGAGCT TCCACATCCG ACCCCTGGTC CCATGCGCTC
AGAGGCCATC GCGCCCTCCA AGGTGTAGC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCGTCGG CAGCTCCTTG CTCTAACAG TGGAGCCAG
TCGCCGAGTA CCAGCGAGCC GTCCAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
TGAATCCGTG TCGTGTACG GGTGTGTGTG CTCACACGGC GTCTTCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAATGAGC GTGGAGATTG GGCTCGCAG
ACCGCCATCC CATACACAGA CTTTACTCG CACCTCTAAC CCGAGCGTGC

1651 CTTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGCCAG
CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTTCTTC TACGTCCGTC

1701 CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGGGTGC
GACTCAACAA CATAAGACTA TTCTAGTCT CCATTGAGG CAACGCCAG

HpaI

1751 TGTTAACGGT GGAGGCCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
 ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NcoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
 GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGCTA

SalINcoIPstIPmlIBclIEcoRV

1851 GGGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA CATATCGCGG
 CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

EcoRI

1901 CCGCCCGGCC GCTCTAGAAT TCTCTAATCA CAGTCATCAT GCGAGCGTCA
 GCGCGCGCCG CGAGATCTTA AGAGATTAGT GTCAGTAGTA CCCTCGCAGT

1951 GGGATTCTGC AATGCCCCG TGAGCGCTTC ACGAAAACAT CTTTCTTGT
 CCTTAAGACG TTAACGGGGC ACTCGCGAAG TCCTTTTGTA GAAAGAAACA

2001 TTGGGTAACTA ATCCTATTCC ATAAAGTCTT TTCAATCCCG TTGGGGGTTG
 AACCCATTAT TAGGATAAGG TATTTCAGAA AAGTTAGGGC AACCCCAAC

2051 TACACAACAA TACCCTACAA GTGAGTGATA TTGACAAGTT TGTGTGCCGA
 ATGTGTCTTT ATGGGATCTT CACTCACTAT AACTGTTCAA ACACACGGCT

2101 GACAACTCTT CTTCAACTAG CCAATTGAAG TCAGTCGGGT TGAAGTTGGA
 CTGTTTGAGA GAAGTTGATC GGTTAACTTC AGTCAGCCCA ACTTGAACCT

2151 GGGCAATGGA CTAGCAACTG ATGTACCAAC GCGAACCAA AGATCGGGTT
 CCCGTTACCT CATCGTTGAC TACATGGTTG CCGTTGGTTT TCTACCCCAA

2201 TTCGAGCTGG TGTCCACCA AAGGTGGTAA ATTACGAAGC TGGAGAATGG
 AAGCTCGACC ACAAGGTGGT TTCCACCATT TAATGCTTCG ACCTCTTACC

2251 GCTGAGAACT GTTATAACCT GGCTATAAAG AAAGTTGATG GTAGTGAGTG
 CGACTCTTGA CAATATTGGA CCGATATTTC TTTCAACTAC CATCACTCAC

2301 CCTACCAGAA GCCCCTGAGG GAGTGAGGGA TTTCCCCCGT TGCCGCTATG
 GGATGGTCTT CGGGGACTCC CTCACTCCCT AAAAGGGCCA ACGGCGATAC

2351 TACACAAAGT CTCAGCAACT GGACCATGCC CAGGAGGACT CGCCTTTCAC
 ATGTGTTTCA GAGTCCTTGA CCTGGTACGG GTCCTCCTGA GCGGAAAGTG

2401 AAAGAAGGAG CCTTCTCTCT GTATGACCGA CTCGCATCAA CAATCATTTA
 TTTCTTCCTC GGAAGAAGGA CATACTGGCT GAGCGTAGTT GTTAGTAAAT

2451 TCGGGGTACA ACCTTTCCCG AAGGAGTTAT TGCATTTCTG ATCTTGCCTA
 AGCCCCATGT TGGAAACGGC TTCCTCAATA ACGTAAAGAC TAGAACGGAT

2501 AGCGCGGAAA GGATTTTTC CAGTCTCCTC CATTCATGA GCCTGCCAAC
 TCCGCGCTTT CCTAAAAAGG GTCAGAGGAG GTAACGTACT CGGACGGTTG

BamHI

2551 ATGACCCACGG ATCCCTCCAG TTACTATCAC ACGACAACAA TAAACTACGT
TACTCGTGCC TAGGGAGGTC AATGATAGTG TGCTGTGTGT ATTGATGCA

2601 GGTTCATAAT TTTGGAACCA ACACCACAGA GTTCTGTTC CAAGTCGATC
CCAACTAATA AACCTTGGT TGTGGTGTCT CAAAGACAAG GTTCAGCTAG

XhoI

2651 ATTTGACGTA TGTGCAGTC GAGGCAAGAT TCACACCACA ATTCTTGTG
TAAACTGCAT ACACGTCGAG CTCCTTCTA AGTGTGGTGT TAAGGAACAG

2701 CTCCTAAATG AAACCATCTA CTCGATAAC CGCAGAAGTA ACACAACAGG
GAGGATTAC TTTGGTAGAT GAGACTATTG GCGTCTTCAT TGTGTTCTCC

2751 AAAACTAATC TGGAAATATA ATCCCACTGT TGATACCAGC ATGGGTGAGT
TTTTGATTAG ACCTTTTATT TAGGGTGACA ACTATGGTCC TACCCACTCA

2801 GCGCTTCTG GGAATATAA AAACTTCAC AAAAACCTT TCAAGTGAAG
CCCGAAAGAC CCTTTTATT TTTGAAGTG TTTTGGGAA AGTTCAC TTC

2851 AGTTGTCTTT CGTACCTGTA CCAGAAACCC AGAACAGGT CCTTGACAGG
TCAACAGAAA GCATGGACAT GGTCTTTGGG TCTTGTCCA GGAACGTGCG

2901 ACACGACGG TCTCTCTCC CATCTCCGC CACAACCAG CAGGCGAAGA
TGTGCTCCC AGAGAGGAGG GTAGAGGCGG GTGTGGTGC GTCCGCTTCT

2951 CCACAAAGAA TTGGTTTCAG AGGATTCCAC TCCAGTGGTT CAGATGCAAA
GGTGTCTCTT AACCAAGTC TCCTAAGGTG AGGTACCAA GTCTACGTTT

3001 ACATCAAGGG AAAGGACACA ATGCCAACCA CAGTACGGG TGTACCAACA
TGTACTCCC TTCTCTGTGT TACGGTGGT GTCACTGCC ACATGGTTGT

HclI

3051 ACCACACCT CTCCTTTCC AATCAATGCT CGCAACACTG ATCATACCAA
TGGCTGGGA GAGGTAAAGG TTAGTTACGA GCGTTGTGAC TAGTATGGTT

3101 ATCATTATC GGCCTGGAGG GCGCCCAAGA AGACCACAGC ACCACACAGC
TAGTAAATAG CCGGACCTCC CCGGGGTCT TCTGGTGTG TGGTGTGTCG

3151 CTGCCAAGAC CACCAGCCAA CCAACCAACA GCACAGAATC GACGACACTA
CACGGTCTG GTGGTCGGT GTTGGTTGT CGTGTCTTAG CTGCTGTGAT

3201 ACCCAACAT CAGAGCCCTC CAGTAGAGG ACGGACCAT CCAGCCCCAC
TTGGGTGTA GTCTCGGAG GTCATCTCCG TGCCCTGGTA GGTGGGGTG

3251 GGTCCCAAC ACCACAGAA GGCACGCCA ACTTGGCAAG ACAACCCAA
CCAGGGTTG TGGTGTCTT CCGTCCGGCT TGAACCGTTC TGTGGGGT

3301 CCACACTCCC AGAACAGCAC ACTGCCGCA GTCCCATTC AAGAGCCGTG
GGTGTGAGG TCTTCTCTG TGACGGCGGT CACGGTAAGG TTCTCGGCAC

3351 CACCCGAGC AACTCAGTG ACCTGGCTC CTGACGAACA CAATACGGGG
GTGGGGCTC TTGAGTCACC TGGACCGAAG GACTGCTTGT GTTATGCCCC

BamHI

3401 GGTGACAAAT CTCCTGACAG GATCCACAAG AAAGCGAAGG GATGTCACTC
CCACTGTTTA GAGGACTGTC CTAGGTCTTC TTTCGCTTCC CTACAGTGAG

3451 CCAATACACA ACCCAAATGC AACCCAAACC TGCCTATTG GACAGCCTTG
GCTTATGTGT TGGGTTTACG TTGGGTTTGG ACGTGATAAC CTGTCGGAAC

3501 GATGAGGGTG CTCCCATAGG TTTAGCCTGG ATACCATACT TCGGGCCAGC
CTACTCCCAC GACGGTATCC AAATCGGACC TATGGTATGA AGCCCGGTCC

3551 AGCTGAGGGA ATTTACACTG AAGGCATAAT CGAGAACTAA AATGGATTGA
TCCACTCCCT TAAATGTGAC TTCCGTATTA CCTCTTAGTT TTACCTAACT

3601 TCTGTGGATT GAGGCAGCTG GCCAACGAAA CGACACAAGC TCTTCAATTG
AGACACCTAA CTCGTCGAC CGCTTGCTTT GCTGTGTTCC AGAAGTTAAC

3651 TTCTTAAGGG CAACTACTGA GTTCGCTACA TTCTCTATAC TAAATCGGAA
AAGAATTCCC GTTGATGACT CAACGCATGT AAGAGATATG ATTAGCCTT

3701 AGCAATAGAC TTCTTGCTCC AAAGATGGGG AGGAACATGT CACATTCTAG
TCCTTATCTG AAGAACGAGG TTTCTACCCC TCCTTGATCA GTGTAAGATC

3751 GGCTGTGATG TTGCATTGAA CCCCAGATT GGACCAAAAA TATCACTGAT
CCGCACTAAC AACGTAACCT GGGGTTCTAA CCTGCTTTT ATAGTGAATA

BclI

3801 AAAATCGATC AAATAATCCA TGACTTTGTC GATAATAATC TTCCAAATCA
TTTAACTAG TTTATTAGGT ACTGAAACAG CTATTATTAG AAGGTTTAGT

3851 GAATGATGCG AGCAACTGGT GGACTGGATG GAAACAATGG GTTCCTGCTG
CTTACTACCG TCGTTGACCA CCTGACCTAC CCTGTGTACC CAAGGACGAC

3901 CAATAGGAAT CACAGGAGTA ATCATTCTTA TTATTGCTTT GCTGTGCATT
CTTATCCTTA GTGTCTCAT TACTAACGAT AATAACGAAA CGACACGTAA

EcoRI

3951 TGCAAAATCA TGCTTTGAAC TAATATAGCA TCATACCTTA GAATCTAGA
ACGCTTAAGT ACGAAACTTG ATTATATCGT AGTATGAAAT CTTAAGATCT

NarIKasIBamHI BclII

4001 CCAGCGCGCT GGATCCAGAT CTGCTGTGCC TTCTAGTTGC CAGCCATCTG
CGTCCGCGGA CCTAGGTCTA GACGACACGG AAGATCAACG GTCGGTAGAC

4051 TTGTTTGCCC CTCCCCGGTG CCTTCCTTCA CCTTGGGAGG TGCCACTCCC
AACAAACGGG GAGGGGGCAC GGAAGGAAGT GGGACCTTC ACCGTGAGGG

4101 ACTGTCCTTT CCTAATAAAA TGAGGAAATT GCATCGCATT GTCTGAGTAG
TGACAGGAAA GGATTATTTT ACTCCTTTAA CGTAGCGTAA CAGACTCATC

4151 GTGTCATTCT ATTCTGGGGG GTGGGGTGGG GCAGCACAGC AAGGGGGAGG
CACAGTAAGA TAAGACCCCC CACCCACCCC CGTCGTGTCC TTCCCCCTCC

10-1

4501 GCTATTAAGT GCAGAGGGAG AGAAAATGCC TCCAACATGT CACGAAGTAA
CGATAATTCA CGTCTCCCTC TCTTTTACGG AGGTTGTACA CTCTTCATT

5001 CTGTGTCAC GAACCCCGG TTCAGCCGA CCGCTGCGC TTATCCGGT
GACACACGTG CTTGGGGGG AAGTCGGCT GCGGACGCG AATAGGCCA

5051 ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA
 TGATAGCAGA ACTCAGGTG GGCATTCTG TGCTGAATAG CGGTGACCGT

 5101 GCAGCCACTG GTAACAGGAT TAGCAGACCG AGGTATGTAG GCGGTGCTAC
 CGTCGGTGAC CATTGTCTTA ATCGTCTCGC TCCATACATC CGCCACGATG

 5151 AGAGTTCTTG AAGTGGTGGC CTAACCTACGG CTACACTAGA AGGACAGTAT
 TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT TCCTGTCATA

 5201 TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT
 AACCATAGAC GCGACACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA

 5251 AGCTCTTGAT CCGGCAAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT
 TCGAGAACTA GCGCGTTTGT TTGGTGGCGA CCATCGCCAC CAAAAAACA

 5301 TTGCAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
 AACGTTCTGC GTCTAATGCG CGTCTTTTTC TCCTAGAGTT CTCTAGGAA

 5351 TCATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA
 ACTAGAAAAA ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT

 5401 GGGATTTTGG TCATGAGATT ATCAAAAAGC ATCTTCACCT AGATCCTTTT
 CCCTAAARCC AGTACTCTAA TAGTTTTTCC TAGAAGTGA TCTAGGAAAA

 5451 AAATTAATAA TGAAGTTTAA AATCAATCTA AAGTATATAT GAGTAAACTT
 TTAATTTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA

 5501 GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT CTCACCGATC
 CCAGACTCTC AATGGTTACG AATTAGTCAC TCCGTGGATA GACTCGCTAG

 5551 TGCTATTTC GTTCATCCAT AGTTGCCTGA CTCGGGGGGG GGGGGCGCT
 ACAGATAAAG CAAGTAGSTA TCAACGGACT GAGGCCCCCC CCCCCCGCA

 5601 GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG ACTCATACCA GGCCTGAATC
 CTCAGACGG ACCACTTCTT CCACAACGAC TGAGTATGGT CCGGACTTAG

 5651 GCCCCATCAT CCAGCCAGAA AGTGAGGGAG CCACGGTTGA TGAGAGCTTT
 CGGGGTACTA GGTCCGTCTT TCACTCCCTC GGTGCCAAT ACTCTCGAAA

 5701 GTGTAGCTG GACCACTTGG TGATTTTGAA CTTTGTCTT GCCACGGAAC
 CACATCCAC CTGCTCAACC ACTAAACTT GAAAACGAAA CGGTGCCTTG

 5751 GGTCTGCGTT GTCCGGAAGA TGCGTGAAT GATCCTTCAA CTCAGCAAAA
 CCAGACGCAA CAGCCCTTCT ACGCACTAGA CTAGGAAGTT GAGTCGTTTT

 5801 GTTCGATTTA TTCAACAAAG CCGCCGTCCC GTCAAGTCAG CGTAATGCTC
 CAACCTAAT AAGTTGTTTC GCGCGCAGGG CAGTTCAGTC GCATTACGAG

 5851 TGCCAGTGTT ACAACCAATT AACCAATTCT GATTAGAAAA ACTCATCCAG
 ACGGTCACAA TGTGGTTAA TTGGTTAAGA CTAATCTTTT TGAGTAGCTC

 5901 CATCAATGA AACTGCAATT TATTCATATC AGGATTATCA ATACCATATT
 GTAGTTTACT TTGACGTAA ATAAGTATAG TCCTAATAGT TATGTATAA

 5951 TTGAAAAAG CCGTTTCTGT AATGAAGGAG AAAACTCACC GAGGCAGTTC
 AACTTTTTTC GGCAAGACA TTACTCCTC TTTGAGTGG CTCCGTCAAG

6001 CATAGGATGG CAAGATCCTG GTATCGGTCT GCGATTCCGA CTCGTCCAAC
GTATCCTACC GTTCTAGGAC CATAGCCAGA CCCTAAGGCT GAGCAGCTTG

6051 ATCAATACAA CCTATTAAAT TCCCTCTGTC AAAAATAAGG TTATCAAGTG
TAGTTATGTT GGATAATTAA AGGGGAGCAG TTTTATTCC AATAGTTCAC

HindIII

6101 AGAATCACC ATGAGTGACG ACTGAATCCG GTGAGAATGG CAAAAGCTTA
TCTTTAGTGG TACTCACTGC TGACTTAGGC CACTCTTACC GTTTTCGAAT

6151 TGCATTTCTT TCCAGACTTG TTCACAGGC CAGCCATTAC GCTCGTCATC
ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG GTCGGTAATG CGAGCAGTAG

6201 AAAATCACTC GCATCAACCA AACCGTTATT CATTCTGTAT TCGCCCTGAG
TTTTAGTGAG CGTAGTTCGT TTGGCAATAA GTAAGCACTA ACCCGGACTC

PvuII

6251 CCAGACGAAA TACGGGATCG CTCTTAAAAG GACAATTACA AACAGGAATC
GCTCTGCTTT ATGCGGTAGC GACAATTTTC CTGTTAATGT TTCTCCTTAG

6301 GAATGCAACC GCGCGAGSAA CACTGCCAGC GCATCAACAA TATTTTCACC
CTTACGTTGG CCGCGTCTTT GTGACGGTCG CGTAGTTGTT ATAAAGTGG

6351 TGAATCAGGA TATTCTTCTA ATACCTGGAA TGCTGTTTTC CCGGGGATCG
ACTTAGTCCT ATAAGAASAT TATGGACCTT ACGACAAAAG GGCCCTAGC

6401 CAGTGGTGAG TAACCATGCA TCATCAGGAG TACGGATAAA ATGCTTGATC
GTCACCACTC ATTGGTACGT AGTAGTCCTC ATGCCTATTT TACGAACATC

6451 GTCCGAAGAG GCATAAATTC CGTCAGCCAG TTTAGTCTGA CCTCTCATC
CAGCCTCTTC CGTATTTAAG GCAGTCGGTC AAATCAGACT GGTAGAGTAG

6501 TGTAACATCA TTGGCAACGC TACCTTTGCC ATGTTTCAGA AACAACTCTG
ACATTGTAGT AACCGTTGCG ATGGAAACGG TACAAAGTCT TTGTTGAGAC

ClaI

6551 GCGCATCGGG CTTCCTATAC AATCGATAGA TTGTCCGACC TGATTGCCCG
CCGCTAGCCC GAAGGGTATC TTAGCTATCT AACAGCGTGG ACTAACGGGC

6601 ACATTATCGC GAGCCCATTT ATACCCATAT AATCAGCAT CCATGTTGGA
TGTAAATAGC CTCGGGTAAA TATGGGTATA TTTAGTCTTA GGTACAACCT

XhoI

6651 ATTTAATCGC GGCCTCGAGC AAGACGTTTC CCGTTGAATA TGGCTCATAA
TAAATTAGCG CCGGAGCTCG TTCTGCAAAG GGCAACTTAT ACCGAGTATT

6701 CACCCCTGCT ATTACTOTTT ATGTAAGCAG ACAGTTTAT TGTTCATGAT
GTGGGAACA TAATGACAAA TACATTGCTC TGTCAAATA ACAAGTACTA

DraIII

6751 GATATATTTT TATCTTCTGC AATGTAAAT CAGAGATTTT GAGACACAAC
CTATATAAAA ATAGAACACG TTACATTGTA GTCTCTAAAA CTCTCTCTTG

Draft

6801 GTCCCTTTCC CCCCCCCCCC ATTATTGAAG CATTATCAG GGTATTGTC
CACCGAAAGG GGGGGGGGGG TAATACTTC GTAAATAGTC CCAATAACAG
.....
6851 TCATGAGCCG ATACATATTT GAATGTATTT AGAAAAATAA ACAATAGGG
AGTACTCGCC TATGTATAAA CTTACATAAA TCTTTTATT TGTATATCC
.....
6901 GTCCGCCCCA CATTCCCCCG AAAACTGCCA CCGACGTCT AAGAAACCAT
CAAGGCGCGT GTAAAGCGGC TTTTCACGGT GCACTGCACA TTCTTTGGTA
.....
6951 TATTATCATG ACATTAACTT ATAAAAATAG GCGTATCAG AGGCCCTTC
ATAATAGTAC TCTAATTGGA TATTTTATC CCGATAGTGC TCCGGGAAG
.....
7001 GTC
CAG
.....

pVR 1012-GP(S)

General Description

DNA pVR 1012-GP(S)

Local object

Created: 09/14/98 03:58PM

Last Modification Date: ? (no data)

length: 7073 bp

storage type: Basic

form: Circular

Comments

Restriction Map

Ball: 1 site TGGCCA
 ACCGGT
 BclI: 1 site TCATCA
 ACTAGT
 ClaI: 1 site ATCGAT
 TAGCTA
 DraIII: 1 site CACNNNGTG
 GTGNNVCAC
 HindIII: 1 site AAGCTT
 TTCGAA
 Kasi: 1 site GCGGCC
 CCGCGG
 KpnI: 1 site GGTACC
 CCATGG
 NarI: 1 site GCGGCC
 CCGCGG
 PmlI: 1 site CACGTG
 GTGCAC
 PvuI: 1 site CGATCG
 GCTAGC
 SacII: 1 site CCGCGG
 GCGGCC
 Sall: 1 site GTCGAC
 CAGCTG
 XbaI: 1 site TCTAGA
 AGATCT
 XmnI: 1 site GAAMNNNTTC
 CTTNNNAAG
 NdeI: 2 sites CATATG
 GTATAC
 EcoRV: 3 sites GATATC
 CTATAG
 SphI: 3 sites GCATGC
 CGTACG
 NcoI: 4 sites CCATGG
 GGTACC
 BamHI: 6 sites GGATCC
 CCTAGG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4090 End: 4642

Kan r

Start: 6138 End: 6760 (Complementary)

Misc_feature (2 signals)**CMV enhancer**

Start: 248 End: 885

GP(S)

Start: 1870 End: 4089

Annotations

1 TCCCGCGTTT CCGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCC
 AGCGGCAAA GCCACTACTG CCACTTTGG AGACTGTGTA CGTCGAGGGC

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCC
 CTCGCGCAST GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGGGGC

101 TCAGGCGCGG TCAGCGGGTG TTGCGGGGTG TCGGGGCTGG CTTAACTATG
 AGTCGCGCGC AGTCGCGCAC AACCGCCAC AGCCCGGACC GAATTGATAC

NdeI

151 CGCCATCAGA CCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
 CCGGTAGTCT CGTCTAACAT GACTCTCAG TGGTATACGC CACACTTTAT

BalI

201 CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
 GCGGTGTCTA CGCATTCCTC TTTTATGGCG TAGTCTAACG GATAACCGGT

251 TTGCATACGT TGTATCCATA TCATAATATC TACATTTATA TTGGCTCATG
 AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAATAT AACCGAGTAC

301 TCCACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
 AGGTGTAAAT GCGGGTACAA CTGTAACATA TAACTGATCA ATAATTATCA

351 ATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGGTT
 TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA

401 ACACTAATTA CCGTAAATGG CCGGCTGGC TGACCGCCCA ACGACCCCG
 TGTATTGAAT GCCATTAC GGGCGGACCG ACTGGCGGGT TGCTGGGGC

451 CCCATTACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
 GGGTAATGCG AGTTATTACT GCATACAAGG GTATCATTGC GGTATCCCT

501 CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
 GAAAGTAAC TCCAGTTACC CACCTCATAA ATGCCATTG ACGGCTGAAC

NdeI

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
 CGTCAGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT

601 TGACGGTAAA TGGCCCCCCT GGCATTATGC CCAGTACATG ACCTTATGGG
 ACTGCCATTI ACCGGGCGGA CCGTAATACG GGTATCTAC TGGAAATACC

NcoI

651 ACTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
 TGAAGCATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

NcoI

701 GTGATCGGGT TTTGGCAGTA CATCAATGGG CGTGGATACC GGTTCAGTC
 CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG

751 ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTCTTTT
 TGGCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAA

801 CCCACCAAAA TCAACGGGAC TTTCACAAAT GTCGTAACAA CTCGCCCCCA
CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT

851 TTGACGCAAA TGGGCGGTAG CCGTGACGG TGGGAGGTCT ATATAAGCAG
AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
TCGAGCAAAT CACTTGCCAG TCTAGCGGAC CTCTCCGGTA GGTCCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGACCGAT CCAGCCTCCG CGGCCGGGAA
AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCCGAGGC GCCGGCCCTT

1001 CCGTGCAATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTTACTGTT
GATATCTGAG ATATCCCTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CTTTATGCTA TAGGTGATGG
AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

1201 TATTGGTGAC GATACTTTC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
GTTGATAGAG ATAACCGATA TACGCTTATG AGACAGGAAG TCTCTGACTG

1301 ACCGACTCTG TATTTTACAA GGATGGGGTC CCAATTATTA TTTACAAATT
TGCCCTGAGAC ATAAAAATG CCAACCCAG GGTAAATAAT AAATGTTTAA

1351 CACATATACA ACAACGGCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
GTGTATATGT TGTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT

1401 CCGTGCGATC TCCACCGGAA TCTCGGGTAC GTGTCCGGA CATGGGCTCT
CGCACCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA

1451 TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
AGAGGCCATC GCCGCCTCGA AGGTGTAGCC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAAATC CCACCACCAC CAGTGTGCCG CACAAGGCCG
TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCAGG
ACCGCCATCC CATAACAGCA CTTTTACTCG CACCTCTAAC CCGAGCGTGC

1651 GCTGACCCAG ATGGAAGACT TTAGGCAGCG GCAGAAGAAG ATCCAGGCAG
CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTTCTC TACGTCCGTC

1701 CTGAGTTCTT GTATTCTGAT AAGAGTCAGA GGTAAGTCCC GTTGCCGTGC
GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG

1751 TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGACCA ACGACGGCGC

NeoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA

Sall

NeoI

PmlI

BclI

EcoRV

1851 GCGTCTTTTC TGCACTCACC GTGGTCGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACCTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

SphI

EcoRV

1901 CCGCTCTAGC TAGATGCATG CTCGAGCGGC CGCCAGTGTG ATGGATATCT
GGCGAGATCG ATCTACGTAC GAGCTCGCCG GCGGTCAACAC TACCTATAGA

NeoI

1951 GCAGAACTCT ATCTTCAGGA TCTCGCCATG GAGGGTCTTA GCCTACTCCA
CGTCTTAAGA TAGAAGTCCT AGAGCGGTAC CTCCCAGAT CGGATGAGGT

2001 ATTGCCCAGA GATAAATTC GAAAAAGCTC TTCTTTGTT TGGGTATCA
TAACGGGTCT CTATTAAAG CTTTTTCGAG AAAGAAACA ACCCAGTAGT

2051 TCTTAATTCA AAAGGCCTTI TCCATGCCCT TGGGTGTTGT GACCAACAGC
AGATAAAGT TTTCCGGAAA AGGTACGGAA ACCCACAAAC CTGGTTGTCTG

2101 ACTTTAGAAG TAACAGAGAT TGACCAGCTA GTCTGCAAGG ATCATCTTGC
TGAAATCTTC ATTGTCTCTA ACTGGTCCAT CAGACGTTCC TAGTAGAACG

2151 ATCAACTGAC CAGCTCAAT CAGTTGGTCT CAACCTCGAG GGGAGCGGAG
TAGTTGACTG GTCGACTTTA GTCAACCAGA GTTGGAGCTC CCCTCGCCTC

EcoRV

2201 TATCTACTGA TATCCCATCT GCGACAAAGC GTTGGGGCTT CAGATCTGGT
ATAGATGACT ATAGGGTAGA CGCTGTTTCG CAACCCCGAA GTCTAGACCA

2251 GTCCCTCCCC AAGTGGTCAG CTATCAAGCA GGAGAATGGG CTGAAAATTG
CACGGACGGG TTCACCACTC GATACTTCCT CCTCTTACCC GACTTTTAAC

2301 CTACAACTCTT GAAATAAAGA AACCGGACGG GAGCGAATGC TTACCCCCAC
GATGCTAGAA CTATTATTTCT TTGGCCTGCC CTGCTTACG AATGCGGGTG

2351 CGCCGGATGG TGTCAGAGGC TTTCCAAGGT CCGCTATGT TCACAAAGCC
GCGGCTACC AAGTCTCCG AAAGGTTCCA CGGCGATACA AGTGTTCGG

2401 CAAGGAACCG GGCCTGCCC GGGTGACTAT GCCTTTCACA AGGATGGAGC
GTTCCTTGGC CCGGACGGG CCCACTGATA CGGAAAGTGT TCCTACCTCG

2451 TTCTTCTCTC CATGACAGGC TGGCTCAAC TGTAATTTAC ACAGGAGTCA
AAAGAAGGAG ATACTGTCCG ACCGAAGTTC ACATTAAATC TCTCCTCAGT

2501 ATTTCTCTGA GGGGTAATC GCATTCTTGA TATTGGCTAA ACCAAAGGAA
 TAAAACGACT CCCCCATTAG CGTAAGAAGT ATAACCGATT TGGTTTCCTT

 2551 ACGTTCCTTC AATCACCCCC CATTGAGAG GCAGCAAACT ACACGAAAA
 TCCAAGGAAG TTAGTGGGGG CTAAGCTCTC CGTCGTTTGA TGTGACTTTT

 2601 TACATCAAGT TACTATGCCA CATCTACTT GGACTACGAA ATCGAAAATT
 ATGTAGTTCA ATGATACGGT GTAGGATGAA CCTCATGCTT TAGCTTTTAA

 2651 TTGGTGCTCA ACACTCCACG ACCCTTTTCA AAATTAACAA TAATACTTTT
 AACCACGAGT TGTGAGGTCC TGGGAAAAGT TTTAATTGTT ATTATGAAAA

 2701 CTCTCTCTGG ACAGGCCCCA CAGCCTCAG TTCCTTTTCC AGCTGAATGA
 CAAGAAGACC TGTCCGGGCT GTCCGGAGTC AAGGAAAAGC TCGACTTACT

 2751 TACCATTCAA CTTACCAAC AGTTGAGCAA CACAACGGG AAATAAATT
 ATCGTAAGTT GAAGTGCTT TCAACTCGTT GTGTTGACCC TTTGATTAAT

 2801 GGCACCTAGA TGCTAATATC AATGCTGATA TTGCTGAATC GGCTTTTTGG
 CCTGTGATCT ACGATTATAG TTACGACTAT AACCACCTTAC CCGAAAAACC

 2851 GAAAATAAAA AAATCTCTCC GAACAACCTAC GTGGAGAAGA GCTGTCTTTC
 CTTTATTTT TTTAGAGAGG CTTGTTGATC CACCTCTTCT CGACAGAAAG

 2901 GAACCTTTAT CGCTCAACGA GACAGAAGAC CATGATGCCA CATGCTCGAG
 CTCTGAAATA GCGAGTTGCT CTGCTCTCTG CTACTACGCT GTAGCAGCTC

 2951 AACTACAAAG GGAAGAATCT CCGACCGGGC CACCAGGAAG TATTGGGACC
 TTGATGTTTC CCTTCTTAGA GGCTGGCCCG GTGGTCCCTC ATAAGCCTGC

 3001 TGGTTCCAAA GGATTCCCTT GGGATGGTTT CATTGCACGT ACCAGAAGGG
 ACCAAGGTTT CCTAAGGGGA CCTACCAA GTAACTGTCA TGGTCTTCCC

 3051 GAAACAACAT TGCCCTCTCA GAATTCGACA GAAGGTGGA GAGTAGATGT
 CTTGTTGTA ACGGCAGAGT CTTAAGCTGT CTCCAGCTT CTCATCTACA

 3101 GAATACTCAG GAACTATCA CAGAGACAAC TGCAACAATC ATAGGCACTA
 CTTATGAGTC CTTTGATAGT GTCTCTGTTG ACGTTGTTAG TATCCGTGAT

 3151 ACGGTAACAA CATGCAGATC TCCACCATCG GGACAGGACT GAGCTCCAGC
 TGCCATTGTT GTACGTCTAG AGGTGCTAGC CCTGTCCTGA CTCGAGGTGC

 NcoI

 3201 CAAATCCTGA GTTCTCACC GACCATGGCA CCAAGCCCTG AACTCAGAC
 GTTAGGACT CAAGGAGTGG CTGGTACCGT GGTTCGGGAC TCTGAGTCTG

 3251 CTCCACAACC TACACACCAA AACTACCAGT GATGACCACC GAGGAACCA
 GAGGTGTTGG ATGCTGCTT TCGATGGTCA CTACTGGTGG CTCCTTGGTT

 3301 CAACACCACC GAGAACTCT CTTGGCTCAA CAACAGAAGC ACCCACTCTC
 GTGTGGTGG CTCTTTGAGA GGACCGAGTT GTACTCTTCG TGGGTGAGAG

 3351 ACCACCCAG AGAATATAAC AACACCGTT AAACTGTTT GGGCACAAGA
 TCGTGGGCTC TCTTATATTG TTGTCGCCAA TTTGACAAA CCCGTGTTCT

3401 GTCCACAAGC AACGGTCTAA TAACCTCAAC AGTAACAGGT ATTCTTGCGA
CAGGTGTTCC TTGCCAGATT ATTGAAGTTG TCATTGTCCA TAAGAACCCT

3451 GCCTTGGACT TCGAAAACGC AGCAGAAAGC AAGTTAACAC CAGGCCACCG
CGGAACCTGA AGCTTTTGGC TCGTCTCTCG TCCAATTGTG GCCCCGGTGC

3501 GGTAAATGCA ATCCCAACTT ACACTACTGG ACTGCCAAG AACAAACATAA
CCATTTACGT TAGGGTTGAA TGTGATGACC TGACGTGTTT TTGTTGTATT

BamHI

3551 TGCTGCTGGG ATTGCCTGGA TCCCGTACTT TGCACCGGGT GCAGAAGGCA
ACGACGACCC TAACCGACCT ACGGCATGAA ACCTGGCCCA CGTCTTCCGT

3601 TATACACTGA AGGCCTTATG CACAACCAAA ATGCCTTAGT CTGTGGACTC
ATATGTCAGT TCCCGAATAC GTGTTGGTTT TACGGAATCA CACACCTGAG

3651 AGCAAACTTG CAAATGAAC AACTCAAGCT CTCCAGCTTT TCTTAAGGGC
TCGTGTGAAC GTTACTTTG TTGAGTTGCA GACGTGAAA AGAATCCCG

3701 CACGACGGAG CTGCGGACAT ATACCATACT CAATAGGAAG GCCATAGATT
GTGCTGCCTC GACGCCTGTA TATGGTATCA GTTATCCTTC CGGTATCTAA

BamHI

3751 TCCTTCTGCG ACGATGGGGC GGGACATGTA GSATCCTGGG ACCAGATTGT
AGGAAGACGC TGCTACCCCG CCCTGTACAT CCTAGGACCC TGGTCTAACA

3801 TCCATTGACC CACATGATTG GACCAAAAC ATCACTGATA AATCAACCA
ACGTAACTCG GTGTACTAAC CTGGTTTTTC TAGTGACTAT TTAGTTGGT

3851 AATCATCCAT GATTTCATCG ACAACCCCTT ACCCAATCAG GATAATGATG
TTASTAGSTA CTAAGTAGC TGTGGGAAA TGGGTTAGTC CTATTACTAC

BamHI

3901 ATAATTGGTG GACGGGCTGG AGACASTGGA TCCCTGCAGG AATAGGCATT
THTTAACCAC CTGCCCCGACC TCTGTACACT AGGGACGTCC TTATCCGTAA

3951 ACTGGAATTA TTATTGCAAT CATTGCTCTT CTTTGGCTCT GCAAGCTGCT
TGACCTTAAT AATAACGTTA GTAACGAGAA GAAACCCAGA CGTTCGACGA

BamHI

4001 TTGTTGAATA TCAGAATTCC AGCACTGGCG GCCGTTACTA GTGGATCCGA
AACAACCTAT AGTCTTAAGG TCGTGACCGC CGGCAATGAT CACCTAGGCT

NarI

BamHI

XbaI

KasI

BamHI

4051 GCTCGGATCC AAGCTCTAGA CCAGGCCGCT GGATCCAGAT CTGCTGTGCC
CGACCTTAGG TTCGAGATCT GGTCCGCGGA CCTAGGTCTA GACGACCGG

4101 TICTACTGTC CAGCCATCTG TTGTTTGCCC CTCCCCCGTC CCTTCCTTGA
AAGATCAACC GTCGGTAGAC AACAAACGGG GAGGGGGCAC GGAAGGAAC

4151 CCTTGAAGG TGCCACTCCC ACTGTCTTTT CCTAATAAAA TGAGGAAATT
GGCACCTTCC ACGGTGAGGG TGACAGGAAA GGATTATTTT ACTCCTTTAA

4201 GCATCGCATT CTCTGAGTAG GTGTCTTCT ATTCTGGGGG GTGGGGTGGG
CGTAGCGTAA CAGACTCATC CACAGTAAGA TAAGACCCCC CACCCACCC

SphI

4231 CCAGCAGAGC AAGCGGGAGG ATTGGGAAGA CAATAGCAGG CATGCTGGGG
CGTCGTGTG TTCCCCCTCC TAACCCTTCT GTTATCGTCC GTACGACCCC

KpnI

4301 ATGCGGTGGG CTCTATGGGT ACCCAGGTGC TGAAGAATTG ACCCGGTTC
TACGCCACCC GAGATACCCA TGGGTCCACG ACTTCTTAAC TGGGCCAAGG

4351 TCCTGGGCCA GAAAGAAGCA GGCACATCCC CTCTCTGTG ACACACCCTG
AGGACCCGGT CTTTCTTCGT CCGTGTAGGG GAAGAGACAC TGTGTGGGAC

4401 TCCACGCCCC TGGTTCTTAG TTCCAGCCCC ACTCATAGCA CACTCATAGC
AGGTGCGGGG ACCAAGAATC AAGGTGCGGG TGAGTATCCT GTCAGTATCG

4451 TCAGGAGGCG TCCGCTTCA ATCCACCCG CTAAGTACT TGGAGCGGTC
AGTCCCTCCG AGCGGAAGT TAGGGTGGG GATTTTATGA ACCTCGCCAG

4501 TCTCCCTCCC TCATCAGCCC ACCAAACCAA ACCTAGCCTC CAAGAGTGGG
AGAGGAGGG AGTAGTCGGG TGGTTTGGT TGCATCGGAG GTTCTCACCC

4551 AAGAAATTAA AGCAAGATAG GCTATTAAAT GCAGAGGGAG AGAAATGCC
TTCTTTAATT TCGTTCTATC CGATAATTCA CGTCTCCCTC TCTTTACGG

XmnI

4601 TCCACATGT GAGGAAGTAA TCAGAGAAAT CATAGAATT CTTCGGCTTC
AGGTTGTACA CTCCTTCATT ACTCTTTA GTATCTTAA GAAGCGAAG

4651 CTCGCTCACT GACTCGCTGC GCTCGGTCTG TCGGCTCGG CGAGCGGTAT
GAGCGAGTGA CTGAGCGAGG CGAGCCAGCA AGCCGACGCC CTTGCGCATA

4701 CAGCTCACTC AAAGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC
GTGAGTGAG TTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCTATTG

4751 GCAGGAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA
CGTCCTTCT TGTACACTCG TTTTCCGGTC GTTTTCCGGT CCTTGCCATT

4801 AAAGGCCGCG TTGCTGGCGT TTTCCATAG GCTCCGCCCC CCTGACGAGC
TTTCCGGCCG AACGACCGCA AAAAGGTATC CGAGCGGGG GGACTGCTCG

4851 ATCACAATAA TCGACGCTCA AGTCAGAGGT GCGGAACCC GACAGGACTA
TAGTGTCTTT AGCTGCGAGT TCAGTCTCCA CCGCTTGGG CTGTCTGAT

4901 TAAAGATACC AGCGGTTTCC CCTGGAACG TCCCTCGTGC GCTCTCTGT
ATTCTATGG TCCGCAAGG GGGACCTTCG AGGGAGCAG CGACAGGACA

4951 TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTCTC CCTTCGGGAA
AGGTGGGAC GCGCAATGC CTATGGACAG CCGGAAGAG GGAAGCCCTT

5001 GCGTGGCGCT TTCTCAATGC TCACGCTGTA GGTATCTCAG TTCGGGTAG
CGCACCGCA AAGAGTTACG AGTGGACAT CCATAGAGTC AAGCCACATC

5051 GTCGTTCCCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA
 CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT

 5101 CCGCTGCCCC TTATCCGGTA ACTATCGTCT TGAATCCAAC CCGTAAGAC
 GGCGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG

 5151 ACCACTTATC GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG
 TGCTGAATAG CCGTGACCGT CGTCGGTGAC CATTTGTCTA ATCGTCTCGC

 5201 AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACGG
 TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC

 5251 CTACACTAGA AGGACAGTAT TTGGTATCTG CGCTCTCTG AAGCCAGTTA
 GACTGTATCT TCCTGTCTA AACCATAGAC GCGACACGAC TTCGGTCAAT

 5301 CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAACA AACCACCGCT
 GGAAGCCTTT TTCTCAACCA TCGAGAACTA GCGCGTTTGT TTGGTGGCGA

 5351 GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAA
 CCATCGCCAC CAAAAAACA AACGTTGCTC GTCTAATGCC CGTCTTTTTT

 5401 AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT
 TCTTAGAGTT CTCTAGGAA ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA

 5451 GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG
 CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC

 5501 ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTAA AATCAATCTA
 TAGAAGTGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTACAGT

 5551 AAGTATATAT GAGTAAACTT GGTCTGACAG TTACCAATGC TTAATCAGTG
 TTCATATATA CTCATTTGAA CCAGACTGTC AATGGTTACG AATTAGTCAC

 5601 AGGCACCTAT CTCAGCGATC TGTCTATTTT GTTCATCCAT AGTTGCCCTGA
 TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT

 5651 CTCCGGGGGG GGGGGGCGCT GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG
 GAGGCCCCCC CCCCCGCGA CTCCAGACGG AGCACTTCTT CCACAACGAC

 5701 ACTCATACCA GGCTGAATC GCCCCATCAT CCAGCCAGAA AGTGAGGGAG
 TGAGTATCGT CCGCACTTAG CGGGGTAGTA GGTCCGTCTT TCACTCCCTC

 5751 CCACGGTTGA TGAGAGCTTT GTTGTAGGTG GACCAGTTGG TGATTTTGAA
 GGTGCCAACT ACTCTCGAAA CAACATCCAC CTGGTCAACC ACTAAAACCT

 5801 CTTTGTCTTT GCCACGGAAC GGTCTCCGTT GTCCGGAAGA TCCGTGATCT
 GAAAAAGAAA CCGTGCCTTG CCAGACGCAA CAGCCCTTCT ACGCACTAGA

 5851 GATCCTTCAA CTCAGCAAAA GTTCGATTTA TTCAACAAAG CCGCCCTCCC
 CTAGGAAGTT GAGTCGTTTT CAAGCTAAAT AAGTTGTTTC GCCGGCAGGG

 5901 GTCAAGTCAG CGTAATGCTC TGCCAGTGTT ACAACCAATT AACCAATTCT
 CAGTTCACTC GCATTACGAG ACGGTCACAA TGTGGTTAA TTGGTTAAGA

 5951 GATTAGAAAA ACTCATCGAG CATCAAATGA AACTGCAATT TATTCATATC
 CTAATCTTTT TGAGTAGCTC GTAGTTTACT TTGACGTTAA ATAAGTATAG

6001 AGGATTATCA ATACCATATT TTTGAAAAG CCGTTTCTGT AATGAAGGAG
 TCCTAATAGT TATGGTATAA AAACCTTTTC GCCAAAGACA TTACTTCCTC

6051 AAAACTCACC GAAGCAGTTC CATAGGATGG CAAGATCCTG GTATCGGTCT
 TTTTGAGTGG CTCCGTCAGG GTATCCTACC GTTCTAGGAC CATAGCCAGA

6101 GCGATTCCGA CTCGTCCAAC ATCAATACAA CCTATTAATT TCCCTCTGTC
 CGCTAAGGCT GAGCAGGTTG TAGTTATGT CGATAATTAA AGGCGACCAG

6151 AAAAATAAGG TTATCAAGTG AGAAATCACC ATGAGTGAGC ACTGAATCCG
 TTTTATTCC AATAGTTCAC TCTTAGTGG TACTCACTGC TGACTTAGGC

HindIII

6201 GTGAGATGG CAAAGCTTA TGCATTTCTT TCCAGACTTG TTCAACAGCC
 CACTCTTACC GTTTTCGAAT ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG

6251 CAGCCATTAC GTCGTGATC AAAATCACTC GCATCAACCA AACCGTTATT
 GTCGTAATG CGAGCAGTAG TTTAGTGAG CCTAGTTGGT TTGGCAATAA

PvuI

6301 CATTCGTGAT TCCGCGTGAG CGAGACGAAA TACCGGATCG CTCTTAAAAG
 GTAAGCACTA ACGCGGACTC GCTCTGCTTT ATGCGCTAGC GACAATTTTC

6351 GACAATTACA AACAGGAATC GAATGCAACC GCGCGAGGAA CACTGCCAGC
 CTGTTAATGT TTGTCCTTAG CTTACGTTGG CCGCGTCCCT GTGACGGTCG

6401 GCATCAACAA TATTTTCACC TGAATCAGGA TACTCTTCTA ATACCTGGAA
 CGTAGTTGTT ATAAAAGTGG ACTTAGTCCT ATAAGAAGAT TATGGACCTC

6451 TGCTCTTTTC CCGGGGATCG CAGTGGTGAG TAACCATGCA TCATCAGGAG
 ACGACAAAAG GCGCCCTAGC GTCACCACTC ATTGGTACGT ACTAGTCCCTC

6501 TACCGATAAA ATGCTTGATC GTCGGAAGAG GCATAAATTC CGTCAGCCAG
 ATCCCTATT TACGAAGTAC CAGCCTTCTC CGTATTTAAG GCAGTCGGTC

6551 TTTACTCTGA CCATCTCATC TGTAACATCA TTGGCAACGC TACCTTTGCC
 AATCAGACT CGTAGAGTAG ACATTGTAGT AACCGTTGCG ATCGAAACGG

ClaI

6601 ATGTTTCAGA AACAACTCTG GCGCATCGGG CTTCCTATAC AATCGATAGA
 TACAAAGTCT TTGTTGAGAC CCGGTAGCCC GAAGGGTATG TTAGCTATCT

6651 TTGTCGCACC TGATTGCCCG ACATTATCCG GAGCCCATTT ATACCATAT
 AACAGCGTGG ACTAACGGGC TGTAATAGCG CTCGGGTAAA TATGGGTATA

6701 AAATCAGCAT CCATGTTGGA ATTTAATCGC GGCCTCGAGC AACACGTTTC
 TTAGTCGTA GGTACAACCT TAAATTAGCG CCGGAGCTCG TTCTGCAAAG

6751 CCGTTGAATA TGGCTCATAA CACCCCTTGT ATTACTGTTT ATGTAAGCAG
 GGCACCTTAT ACCGAGTATT GTGGGGAACA TAATGACAAA TACATTGCTC

6801 ACAGTTTAT TGTTATGAT GATATATTTT TATCTTGTC AATGTAACAT
 TGTCAAAATA ACAAGTACTA CTATATAAAA ATAGAACACG TTACATTGTA

DraIII

6851 CAGAGATTTT GAGACACAAC GTGGCTTTCC CCCCCCCCCC ATTATTGAAG
 CTCTCTAAAA CTCGTGTGTC CACCCAAAGG GGGGGGGGGG TAATAACTTC

6901 CATTEATCAG GGTATTGTC TCATGAGCGC ATACATATTT GAATGTATTT
 GTAAATAGTC CCAATAACAG AGTACTCGCC TATGTATATA CTTACATAAA

6951 AAAAAATATA ACAAATAGGG GTTCCGCCCA CATTCCCCCG AAAAGTGCCA
 TCITTTTATT TGTTTATCCC CAAGGCCCGT GTAAAGGGGC TTTCACCGT

7001 CCTGACGTCT AAGAAACCAT TATTATCATC ACATTAACTT ATAAAAATAG
 GCACTGCAGA TTCTTTGGTA ATAATAGTAC TGTAAATTGA TATTTTATC

7051 GCGTATCAG AGGCCCTTTC CTC
 CGCATAGTGC TCCGGGAAAG CAG

SEQUENCE LISTING ID NO: 3

pVA 1012-GP(Z)

General Description

DNA pVR 1012-GP(Z)
 Local object
 Created: 09/15/98 05:06PM
 Last Modification Date: ? (no data)
 length: 7285 bp
 storage type: Basic
 form: Circular

Comments

Restriction Map

DraIII: 1 site CACGCGTG
 GTGCGGCAC

HindIII: 1 site AAGCTT
 TTCGTA

HpaI: 1 site GTTAC
 CAATTG

KasI: 1 site GCGGCC
 CCTCGG

NarI: 1 site GCGGCC
 CCTCGG

NotI: 1 site GCGGCCGC
 CGCCGCGC

PmlI: 1 site CACGTG
 GTGCAC

PvuI: 1 site CGATCG
 GCTAGC

SacII: 1 site CCGCGG
 GCGGCC

XbaI: 1 site TCTAGA
 AGATCT

XhoI: 1 site CTCGAG
 GAGCTC

EcoRV: 2 sites GATATC
 CTATAG

NcoI: 2 sites CCATCG
 GGTACC

NdeI: 2 sites CATATG
 GTATAC

SphI: 2 sites GCATCG
 CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4302 End: 4854

Kan r

Start: 6350 End: 6972 (Complementary)

Misc_feature (2 signals)**CMV enhancer**

Start: 248 End: 885

GP(Z)

Start: 1870 End: 4301

Annotations

T04T50" 99/00950

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
ACCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

51 GAGACGGTCA CAGCTTGTCT GTAACCGGAT GCCGGGAGCA GACAAGCCCC
CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGGGGC

101 TCAGGGCGCG TCACCGGGTG TTGCGGGGTG TCGGGGCTCG CTTAACTATG
AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCGGACC GAATTGATAC

NdeI

151 CGGCATCACA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
CCCGTAGTCT CGTCTAACAT GACTCTCAGG TGGTATACCG CACACTTTAT

201 CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
GGCGTGCTA CGCATTCCTC TTTTATGCGG TAGTCTAACC GATAACCGGT

251 TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
AACGTATCCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

301 TCCAACATTA CGGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
AGGTTGTAAT GCGGGTACAA CTGTAACATA TAACTGATCA ATAAATTATCA

351 AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
TTAGTTAATG CCCAGTAAT CAAGTATCGG GTATATACCT CAAGCGGCAA

401 ACATAACTTA CCGTAAATGG CCGCGCTGGC TCACCGCCCA ACGACCCCG
TGTATGAAT CCCATTATCC GGGCGGACCG ACTGGCGCGT TGCTGGGGGC

451 CCCATTGACG TCAATAATGA CGTATGTCC CATAGTAACG CCAATAGGGA
GGGTAAGTGC AGTTATFACT GCATACAAGG GTATCATGCG GGTATCCCT

501 CTTTCCATTG ACGTCAATGG GTGGAGTAT TACGGTAAAC TGCCCACTTG
GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTG ACGGGTGAAC

NdeI

551 GCAATACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
CGTCATGTAG TTCACATAGT ATACGGTTCA TCGGGGGGAT AACTGCAGTT

601 TGACCGTAAA TGGCCCCCCT GGCATTATGC CCAGTACATG ACCTTATGGG
ACTGCCATTG ACCGGCGCGA CCGTAATACG GGTCAATGAC TGGAAATACCC

NcoI

651 ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
TGAAACGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

NcoI

701 GTGATCGCGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG

751 ACGGGGACTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTGTT
TGCCCCATAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAAATAAA

801 GCCACCAAAA TCAACGGGAC TTCCAAAAAT GTCGTAAACA CTCCGCCCCA
CCGTGCTTTT AGTTGCCCTC AAAGGTTTAA CAGCATTTGT GAGGCGGGGT

851 TTGACGCAAA TGGCGCGTAG GCGTGACGG TCGGAGGTCT ATATAAGCAG
AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTGTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACCCAT CCACCGTGT
TCGACCAAT CACTTGGCAG TCTAGCGGAC CTCTCGGTA GGTGCGACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTGGGAGG CCGGCCCTT

1001 CCGTGCAATG GAACCGCGAT TCCCGTGCC AAGAGTCAG TAAGTACCGC
GCCACGTAAC CTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTAFACTGT
GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG CGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGIGATGG
AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTATTG ACCATTATTG ACCACTCCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

1201 TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCTTC AGAGACTGAC
GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTACA GGATGGGGTC CCATTTATTA TTTACAAAT
TGCTGAGAC ATAAAAATGT CCTACCCAG GGTAAATAT AATGTTTAA

1351 CACATATACA ACAACGCCGT CCCCCGTCC CGCAGTTTT ATTAAACATA
GTGTATATGT TGTGCGGCA GGGGGCACGG CCGTCAAAA TAATTTGTAT

1401 GCGTGGGATC TCCACCGGAA TCTCGGGTAC GTGTCCGGA CATGGGCTCT
CGCACCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA

1451 TCTCCGGTAG CGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
AGAGGCCATC CCGCCCTCGA AGGTGTAGG TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCTAACAG TGGAGGCCAG
TCCCGAGTA CCAGCCAGCC GTCGAGGAAC GAGGATTCTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACAAATC CCACCACCAC CAGTGTCCG CACAAGCCG
TGAATCCGTG TCGTGTACG GGTGGTGGT GTCACACGGC GTGTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAATGAGC GTGCACATTG GGCTCGCAG
ACCGCCATCC CATAACAGA CTTTACTCG CACCTCTAAC CCGAGCGTC

1651 GCTGACGCAG ATGGAAGACT TAAGCCAGCG CCAGAAGAAG ATGCAGGCAG
CGACTGCGTC TACCTTCTGA ATTCCGTCCG CGTCTTCTTC TACGTCCGTC

1701 CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGGGTGC
GACTCAACA CATAAGACTA TTCTAGTCT CCATTGAGGG CAACGCCACG

HpaI

1751 TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TCCTGCCGCG
ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NcoI

1801 CCGGCCACCA GACATAATAG CTCACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAACGTA

NcoIPmlIEcoRVNotI

1851 GGGTCTTTTC TGCAGTCACC GTCCGCGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT CCACACTAGT CTATAGCGCC

NarINotI XbaIKasI

1901 CCGCTCTAGA CCAGGCGCCT GGATCGATCC GCGATGAAGA TTAAGCCGAC
GGCGAGATCT CGTCCGCGCA CCTAGCTAGG CGCTACTTCT AATTCCGGCTG

1951 AGTCAGCGTA ATCTTCATCT CTCITAGATT ATTTGTTTTC CAGAGTAGGG
TCACTCGCAT TAGAAGTAGA GAGAATCTAA TAAACAAAAG GTCTCATCCC

2001 GTCGTCAGGT CTTTTTCAAT CGTGTAACCA AAATAAATC CACTAGAAGG
CAGCAGTCCA GCAAAAGTTA GCACATTGGT TTTATTGAG GTGATCTTC

2051 ATATTGIGGG GCAACAACAC AATGGCGGTT ACAGGAATAT TCCAGTTACC
TATAACACCC CGTTGTTGTG TTACCCGCAA TGTCCTTATA ACGTCAATGG

2101 TCCTGATCGA TTCAAGAGGA CATCATCTTT TCTTTGGGTA ATTATCCTTT
AGCACTAGCT AAGTTCICCT GTAGTAAGAA AGAAACCCAT TAATAGGAAA

2151 TCCAAAGAAC ATTTTCCATC CCACTTGGAG TCATCCACAA TAGCACATA
AGGTTTCTTG TAAAGGTAG GGTGAACCTC ACTAGGTGTT ATCGTGTAAT

2201 CAGGTTAGTG ATGTCCGACAA ACTAGTTTGT CGTGACAAAC TGTCATCCAC
GTCCAATCAC TACAGCTGTT TGATCAAAAC GCACCTTTTG ACAGTAGGTG

2251 AATCAATTG AGATCAGTTG GACTGAATCT CGAAGCGAAT GGAGTGGCAA
TTTAGTTAAC TCTAGTCAAC CTGACTTAGA GCTTCCCTTA CCTCACCGTT

2301 CTGACGTGCC ATCTGCAACT AAAACATGGG GCTTCAGGTC CGGTGTCCCA
GACTGCACGG TAGACGTTGA TTTCTTACCC CGAAGTCCAG GCCACAGGGT

2351 CCAAGGTTGG TCAATTATGA ACCTGGTGAA TGGGCTGAAA ACTGCTACAA
GGTTTCCACC AGTTAATACT TCGACCACTT ACCCGACTTT TGACGATGTT

2401 TCTTGAAATC AAAAAACCTG ACGGCGAGTGA GTGTCTACCA GCAGCGCCAG
AGAACTTTAG TTTTTTGGAC TGCCCTCACT CACAGATGCT CGTCGCGGTC

2451 ACGGGATTCTG GGGCTTCCCC CGGTGCCCGT ATGTCCACAA ACTATCAGGA
TGCCCTAAGC CCCGAAGGGG GCCACGGCCA TACACGTGTT TCATAGTCCT

2501 ACGGGACCGT GTGCCCGAGA CTTTGCCCTTC CATAAAGAGG GTGCTTCTTT
TGCCCTGGCA CACGGCCTCT GAAACGGAAG GTATTTCTCC CACGAAGAA

2551 CCTGTATGAT CGACTTGCTT CCACAGTTAT CFACCGAGGA ACGACTTTTCG
GGACATCTA GCTGAACGAA GGTGTCAATA GATGGCTCCT TCGTGAAAGC

2601 CTGAAGGTGT CGTTGCATTT CTGATACTGC CCCAAGCTAA GAAGGACTTC
GACTTCCACA GCAACGTAAA GACTATGACC GGGTTCGATT CTTCCTGAAG

2651 TTCAGCTCAC ACCCCTTGAG AGAGCCGGTC AATGCAACGG AGGACCCGTC
AAGTCGAGTG TGGCGAATC TCTCGGCCAG TTACGTTGCC TCCTCGGCAG

EcoRV

2701 TAGTGGCTAC TATTCTACCA CAATTAGATA TCAGGCTACC GGTTTTGGAA
ATCACCGATG ATAAGATGGT GTTAATCTAT AGTCCGATCG CCAAAACCTT

2751 CCAATGAGAC AGAGTACTTG TTCGAGGTTG ACAATTTGAC CTACGTCCAA
GGTFACTCTG TCTCATGAAC AAGCTCCAAC TGTAAACTG GATCCAGGTT

2801 CTTGAATCAA GATTACACAC ACAGTTTCTG CTCCAGCTGA ATGAGACAAT
GAACCTAGTT CTAAGTGTGG TGTCAAAGAC GAGGTCGACT TACTCTGTTA

2851 ATATACAAGT GGGAAAAGCA GCATACCAC GCGAAAATA ATTGGAAGG
TATATGTTCA CCCTTTTCCT CGTTATGGTG CCCTTTTGAT TAAACCTTCC

2901 TCAACCCCGA AATTGATACA ACAATCGGGG AGTGGGCTT CTGGGAACCT
AGTTGGGGCT TTAATATGTT TGTAGCCCC TCACCCGGA GACCCTTTGA

2951 AAAAAAACC TCACTAGAAA AATTCGCAGT GAAGAGTTGT CTTTACAGT
TTTTTTTTGG AGTGATCTTT TTAAGCGTCA CTCTCAACA GAAAGTGTC

3001 TGTATCAAC GGAGCCAAAA ACAICAGTGG TCAGAGTCCG GCGCGAAGTT
ACATAGTTTG CCTCGGTTTT TGTAGTCACC AGTCTCAGGC CGCGCTTGAA

3051 CTTCCGACCC AGGGACCAAC ACAACAAGTG AAGACCACAA AATCATGGCT
GAAGGCTGGG TCCCTGGTTG TGTGTGTGAC TTCTGGTGT TTAGTACCGA

3101 TCAGAAAATT CCTCTCCAA GGTTCAGTG CACAGTCAAG GAAGCGAAGC
AGTCTTTTAA GGAGACGTTA CCAAGTTCAC GTGTCAAGTC CTTCCTTCG

3151 TCCAGTGTGG CATCTAACAA CCCTTGCCAC AATCTCCACG AGTCCCAAT
ACGTCACAGC GTAGATTGTT GGGAACGGTG TTAGAGGTGC TCAGGGCTTA

3201 CCCTCACAAC CAAACCAAGT CCGGACACAA GCACCCATAA TACACCGTG
GGGAGTGTTC GTTTCCTCCA GGCCTGTTGT CGTGGGTATT ATGTGGGCAC

3251 TATAAACTTG ACATCTCTGA GGCAACTCAA GTTGAACAC ATCACCGCAG
ATATTGAAAC TGTAGAGACT CCGTTCAGTT CAACCTGTTC TAGTGGCCTC

3301 AACAGACAAC GACAGCACAG CCTCCGACAC TCCCTCTGCC ACGACCGCAG
TTGTCTGTTG CTCTCGTGTG GGAGGCTGTG AGGGAGACGG TCCTGGCCTC

3351 CCGGACCCCC AAAAGCAGAG AACACCAACA CGAGCAAGAG CACTGACTTC
GGCCTGGGGG TTTCCTCTC TTGTGTTGT GCTCGTTCTC GTGACTGAAG

3401 CTGGACCCCG CCACCACAAC AAGTCCCCAA AACACAGCG AGACCGCTGG
GACCTGGGCG GGTGGTGTG TTCAGGGGTT TTGGTGTCCG TCTGGCGACC

3451 CAACAACAAC ACTCATCACC AAGATACCGG AGAAGAGAGT GCCAGCAGCG
 GTTGTGTTG TGAAGTAGTG TTCTATGGCC TCTTCTCTCA CGGTGCTCGC

 3501 GGAAGCTAGG CTTAATTACC AATACTATTG CTGGAGTCCG ACGACTGATC
 CCTTCGATCC GAATTAATGG TTATGATAAC GACCTCAGCG TCCTGACTAG

 3551 ACACGCGGGA GAAGAACTCG AAGAGAAGCA ATTGTCAATG CTCACCCCAA
 TGTCCGCCCT CTTCTTGAGC TTCTCTTCGT TAACAGTTAC GAGTTGGGTT

 3601 ATGCAACCCCT AATTACATT ACTGGACTAC TCAGGATGAA GGTGCTGCAA
 TACGTTGGGA TTAATGTAA TGACCTGATG AGTCCTACTT CCACGACGTT

 3651 TCGACTGGC CTGGATACCA TATTTCGGGC CAGCAGCCGA GCGAATTTAC
 AGCCTGACCG GACCTATGGT ATAAAGCCCG GTCGTGCGCT CCCTTAAATC

 3701 ATAGAGGGGC TAATGCACAA TCAAGATGGT TTAATCTGTG GGTGAGACA
 TATCTCCCGG ATTACGTGTT AGTTCTACCA AATTAGACAC CCAACTCTGT

 3751 GCTGGCCAAAC GAGACGACTC AAGCTCTTCA ACTGTTCTTG AGAGCCACAA
 CGACCGGTG CTCTGCTGAG TTCGAGAAGT TGACAAGGAC TCTCGGTGTT

 3801 CTGAGCTACG CACCTTTTCA ATCCTCAACC GTAAGGCAAT TGATTCTTG
 GACTCGATGC GTGGAAAAGT TAGGAGTTGG CATCCGTTA ACTAAAGAAC

 3851 CTGCAGCGAT GGGCGGGCAC ATGCCACATT CTGGGACCGG ACTGCTGTAT
 GACGTGCGTA CCCCGCCGTG TACGGTGTA GACCCTGGCC TGACGACATA

 3901 CGAACCACAT GATTGCACCA AGAACATAAC AGACAAAAT GATCAGATTA
 GCTTGCTGTA CTAACCIGGT TCTTGATTG TCTGTTTAA CTAGTCTAAT

 3951 TTCAATGATT TGTGATAAA ACCCTTCCGG ACCAGGGGGA CAATGACAA
 AAGTAATAA ACAACTATTT TGGGAAGGCC TGCTCCCCCT GTTACTGTTA

 4001 TGGTGGACAG GATGGAGACA ATGGATACCG GCAGGTATTG GAGTTACAGG
 ACCACCTGTC CTACCTCTGT TACCTATGGC COTCCATAAC CTCAATGTCC

 4051 CGTTATAATT GCAGTTATCG CTTTATTCTG TATATGAAA TTGTCTTTT
 GCAATATTAA CGTCAATAGC GAAATAAGAC ATATACGTTT AAACAGAAAA

 4101 AGTTTTCTTT CAGATTGCTT CATGCAAAAG CTCAGCCTCA AATCAATGAA
 TCAAAAAGAA GTCTAACGAA GTACCTTTTC GAGTCGGAGT TTAGTTACTT

 4151 ACCAGGATTT AATTATATGG ATTACTTGAA TCTAAGATTA CTTGACAAAT
 TGGTCCTAAA TTAATATACC TAATGAACCT AGATTCTAAT GAAGTGTTTA

 4201 GATAATATAA TACACTGGAG CTTTAAACAT AGCCAATGTG ATTCTAACTC
 CTATTATATT ATGTGACCTC GAAATTGTA TCGGTTACAC TAAGATTGAG

 4251 CTTTAAACTC ACAGTTAATC ATAAACAAGG TTTGGTACCG AGCTCGAATT
 GAAATTTGAG TGTCAATAG TATTTGTTCC AAACCATGCC TCGAGCTTAA

 4301 ATCTGCTGTG CCTTCTAGTT GCCAGCCATC TGTGTTTGC CCTCCCCCG
 TAGACGACAC GGAAGATCAA CGGTGCGTAG ACAACAAACG GGGAGGGGGC

 4351 TGCCTTCTTT GACCTGGA GGTGCCACTC CCACTGTCTT TTCCTAATAA
 ACGGAAGGAA CTCGGACCTT CCACGGTGAG GGTGACAGCA AAGGATTATT

4401 AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG
TTACTCCTTT AACGTAGCGT AACAGACTCA TCCACAGTAA GATAAGACCC

4451 GGGTGGGGTG GGGCAGGCACA GCAAGGGGGA GGATTGGGAA GACAATACCA
CCCACCCAC CCCGTCTGTG GTTCCCCCTT CCTAACCCCT CTGTTATCGT

SphI

4501 GGCATGCTGG GGATGCGGTG GGTCTATGCG GTACCCAGGT GCTGAAGAAT
CCGTACGACC CCTACGCCAC CCGAGATACC CATGGGTCCA CGACTTCTTA

4551 TGACCCGGTT CCTCCTGGGC CAGAAAGAAG CAGGCACATC CCCTTCTCTG
ACTCGGCCAA GGAGGACCCG GTCTTTCTTC GTCCGTGTAG GGGAGAGAG

4601 TGACACACCC TGTCCACGCC CCTGGTTCTT AGTTCACGCC CCACTCATAG
ACTGTGTGGG ACAGGTGCGG GGACCAAGAA TCAAGGTCCG GGTGAGTATC

4651 GACACTCATA GCTCAGGAGG GCTCCGCCTT CAATCCCACC CGCTAAAGTA
CTGTGAGTAT CGAGTCCTCC CGAGCGCGAA GTTAGGGTGG CCGATTTCAT

4701 CTGGGAGCGG TCTCTCCCTC CCTCATCAGC CCACCAAACC AAACCTAGCC
GAACCTCGCC AGAGAGGGAG CGAGTAGTCC GGTGTTTGG TTTGGATCGG

4751 TCCAAGAGTG GGAAGAAATT AAAGCAAGAT AGGCTATTAA GTGCAGAGGG
AGGTTCTCAC CCTTCCTTAA TTCTGTTCTA TCCGATAATT CACGTCTCCC

4801 AGAGAAAATG CCTCCAACAT GTGAGGAAGT AATGAGAGAA ATCATAGAAT
TCTCTTTTAC GGAGGTTCTA CACTCCTTCA TTAATCTCTT TAGTATCTTA

4851 TTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC
AAGAAGGCCA AGGAGCCGAGT GACTGAGCGA CCGGAGCCAG CAAGCCGACG

4901 GGCAGCGGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA
CCCTCGGCA TAGTCGAGTG AGTTTCCGCC ATATATGCAA TAGGTGTCTT

4951 TCAGGGGATA ACCGAGGAAA GAACATGTGA GCAAAAGCCC AGCAAAAGGC
AGTCCCTAT TCGCTCCTTT CTGTGACACT CGTTTTCCGG TCGTTTTCCG

5001 CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC
GTCCTTGGCA TTTTTCGGCG GCAACGACCG CAAAAAGGTA TCCGAGGCGG

5051 CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAGC
GGGACTGCT CGTAGTGTTT TTAGCTCCGA GTTCAGTCTC CACCGCTTTG

5101 CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCTGGAA GCTCCCTCGT
GGCTGTCTCT ATATTCTAT GGTCCGCAAA GCGGGACCTT CGAGGGAGCA

5151 GCGCTCTCTT GTTCCGAGCC TGCCGCTTAC CGGATACCTG TCCGCTTTTC
CGCGAGAGGA CAAGGCTGGG ACGGCGAATG GCCTATGGAC AGGCGGAAAG

5201 TCCCTTCGGG AAGCOTGGCG CTCTCTCAAT GCTCACGCTG TAGGTATCTC
AAGGAAGCCC TTCGCACCGC GAAAGAGTTA CGAGTGCGAC ATCCATAGAG

5251 AGTTCGCTGT AGGTCTGTCG CTCCAAGCTG GCGTGTGTGC ACGAACCCCG
TCAAGCCACA TCCAGCAAGC GAGGTTCGAC CCGACACACG TGCTTGGGGG

5301 CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA
 GCAAATCGGG CTGGCGACGC GGAATAGGCC ATTGATAGCA GAACTCAGGT

 5351 ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG
 TCGGCCATTG TGTGCTGAAT AGCGGTGACC GTCGTCCGTG ACCATTGTTC

 5401 ATTAGCAGAG CGAGGTATGT AGCGGTGCT ACAGAGTTCT TGAAGTGGTG
 TAACTCGTCT GCTCCATACA TCCGCCACGA TGTCTCAAGA ACTTCACCAC

 5451 GCCTAACTAC GGCTACACTA CAAGGACAGT ATTTGGTATC TGGCCTCTGC
 CGGATTGATG CCGATGTGAT CTTCTGTCA TAAACCATAG ACCCGAGACG

 5501 TGAAGCCAGT TACCTTCGCA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA
 ACTTCGGTCA ATGGAAGCCT TTTTCTCAAC CATCGAGAAC TAGGCCGTTT

 5551 CAAACCAACG CTGGTAGCGG TGGTTTTTTT GTTTCGAAGC AGCAGATTAC
 GTTGTGTGCC GACCATCGCC ACCAAAAAAA CAAACGTTCC TCGTCTAATG

 5601 CCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT
 CGGCTCTTTT TTCTCTAGAG TTCTTCTAGG AAAGTAGAAA AGATGCCCCA

 5651 CTGACCGTCA GTGGAACGAA AACTCAGGTT AAGGGATTTT GGTCATGAGA
 GACTGCGAGT CACCTTGCTT TTGAGTGCAA TTCCCTAAAA CCAGTACTCT

 5701 TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAATTAATA AATGAAGTTT
 AATAGTTTTT CCTAGAAGTG GATCTAGCAA AATTTAATTT TACTTCAAA

 5751 TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT
 ATTAGTTAG ATTTCAATATA TACTCATTG AACCAAGCTG TCAATGGTTA

 5801 GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC
 CGAATTAGTC ACTCCGTGGA TAGAGTCCCT AGACAGATAA AGCAAGTAGG

 5851 ATAGTTGCCCT CACTCCGCGG GGGGGGGGCG CTGAGGCTG CCTCGTGAAG
 TATCAACCGA CTGAGGCCCG CCCCCCGCGC GACTCCAGAC GGAGCACTTC

 5901 AAGGTGTTGC TGAATCATAC CAGGCCTGAA TCGCCCCATC ATCCAGCCAG
 TTCCACAACG ACTGAGTATG GTCCGGACTT AGCGGGGTAG TAGGTCGGTC

 5951 AAAGTGAGGG AGCCACGGTT GATGAGAGCT TTGTTGTAGG TGCACCAATT
 TTCACTCCC TCGGTGCCAA CTAATCTCGA AACAAACATCC ACCTGGTCAA

 6001 GGTGATTTTG AACTTTTGCT TTGCCACGGA ACGGTCTGCG TTGTCGGGAA
 CCACTAAATC TTGAAAACGA AACGGTGCCT TGCCAGACCG AACAGCCCTT

 6051 GATCGGTGAT CTGATCCTTC AACTCAGCAA AAGTTGCAAT TATTCAACAA
 CTACGCACTA GACTAGGAAG TTGAGTCCGT TTCAAGCTAA ATAAGTTGTT

 6101 AGCCGCGGTC CCGTCAAGTC AGCGTAATGC TCTGCCAGTG TTACAACCAA
 TCGCGGOCAG CGCAGTTCAG TCGCATACG AGACGGTCAC AATGTTGGTT

 6151 TTAACCAATT CTGATTAGAA AAATCATCG AGCATCAAT GAAACTGCAA
 AATTGGTTAA GACTAATCTT TTTGAGTAGC TCGTAGTTTA CTTTGACGTT

 6201 TTTATTCATA TCAGGATTAT CAATACCATA TTTTGAAGAA AGCCGTTTCT
 AAATAAGTAT AGTCTTAATA GTTATGGTAT AAAAATTTTT TCGGCAAGA

6251 GTAATGAAGG AGAAAACCTCA CCGAGGCAGT TCCATAGGAT GGCAAGATCC
CAATACCTTC TCTTTTGAGT GGCTCCCTCA AGGTATCCTA CCGTTCTAGG

6301 TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC AACCTATTAA
ACCATAGCCA GACGCTAAGG CTGAGCAGGT TGTAGTTATG TTGGATAATT

6351 TTTCCCTCTG TCAAAAATAA GGTATCAAG TGAGAAATCA CCATGAGTGA
AAAGGGGACG AGTTTTTATT CCAATAGTTC ACTCTTTAGT GGTACTCACT

HindIII

6401 CGACTGAATC CGGTGACAAT GGCAAAAGCT TATGCATTC TTTCCAGACT
GCTGACTTAG GCCACTCTTA CCGTTTTTCA ATACGTAAAG AAAGGTCTGA

6451 TGTTC AACAC GCGAGCCATT ACGCTCGTCA TCAAAATCAC TCGCATCAAC
ACAAGTTGTC CCGTCCGGTAA TCGGAGCACT AGTTTTAGTG AGCGTAGTTG

PvuI

6501 CAAACCGCTA TTCATTCTGT ATTGCGCCTG AGCGAGACGA AATACGCGAT
GTTTGGCAAT AAGTAAGCAC TAACCGCGAC TCGCTCTGCT TTATGCGCTA

PvuI

6551 CGCTGTAAAA AGGACAATTA CAAACAGGAA TCGAATGCAA CCGCGCCAGG
CCGACAATTT TCCTGTAAAT GTTTGTCTTT AGCTTACGTT GGCCGCGTCC

6601 AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG GATACTCTTC
TTGTACGGT CCGGTAGTTG TTATAAAGT GCACTTAGTC CTATAAGAAG

6651 TAATACCTGG AATGCTGTTT TCCCGGGGAT CGCACTGGTG AGTAACCATG
ATTATGGACC TTACGACAAA AGGGCCCTTA GCGTCACCAC TCATTGGTAC

6701 CATCATCAGG AGTACGGATA AAATGCTTGA TGGTCGGAAG AGGCATAAAT
GTAGTAGTCC TCATGCCAT TTTACGAACT ACCAGCCTTC TCCGTATTTA

6751 TCCGTCAGCC AGTTTAGTCT GACCATCTCA TCTGTAAAT CATTGGCAAC
AGGCAGTCGG TCAAAATCAGA CTGGTAGAGT AGACATTGTA GTAACCGTTG

6801 CCTACCTTTG CCATGTTTCA GAAACAACCT TGGCGCATCG GGCTTCCCAT
CGATGGAAAC GGTACAAAGT CTTTGTGAG ACCGCGTAGC CCGAAGGGTA

6851 ACAATCGATA GATTCTCGCA CCTGATTGCC CGACATTATC CCGAGCCCAT
TGTTAGCTAT CTAACAGCGT GGACTAACGG CCGTGAATAG CGCTCGGGTA

XhoI

6901 TTATACCCAT ATAAATCACC ATCCATGTTG GAATTTAATC GCGGCCCTCA
AATATGGGTA TATTTAGTGG TAGGTACAAC CTTAAATTAG CGCCGGAGCT

XhoI

6951 GCAAGACGTT TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT
CGTTCTGCAA AGGGCAACTT ATACCGAGTA TTGTGGGCAA CATAATCACA

7001 TTATGTAAGC AGACAGTTTT ATTGTTTCATG ATGATATATT TTTATCTTGT
AATACATTCTG TCTGTCAAAA TAACAAGTAC TACTATATAA AAATAGAACA

DraIII

7051 GCAATGTAAC ATCAGAGATT TTGAGACACA ACGTGGCTTT CCCCCCCCCC
CGTTACATTG TAGTCTCTAA AACTCTGTGT TGCACCGAAA GGGGGGGGGG

7101 CCATTATTGA AGCATTATC AGCGTCATTG TCTCATGAGC GCATACATAT
GGTAATAACT TCGTAAATAG TCCCAATAAC AGAGTACTCC CCTATGTATA

7151 TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCGCGC CACATTTCCT
AACTTACATA AATCTTTTAA TTGTTTATC CCCAAGGCGC GTGTAAAGGG

7201 CCAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAAC
GCTTTTCACG GTGGACTGCA GATTCTTTGG TAATAATAGT ACTGTAATTG

7251 CTATAAAAAT AGGCGTATCA CGAGGCCCTT TCGTC
GATATTTTAA TCCGCATAGT GCTCCGGGAA AGCAG

09200960
054401

pVR 1012-SCP(Z)

General Description

DNA pVR 1012-SCP(Z)

Local object

Created: 09/14/98 04:29PM

Last Modified: 09/15/98 04:50PM

length: 7272 bp

storage type: Basic

form: Circular

Comments

Restriction Map

DraIII: 1 site CACNNGGTG
 GTGNWNCAC

HindIII: 1 site AAGCTT
 TTCGTA

HpaI: 1 site GTTAAC
 CAATTG

KpnI: 1 site GGTACC
 CCATGG

NotI: 1 site GCGGCCCG
 CGCCGGCG

PmlI: 1 site CACGTC
 GTGCAC

PvuI: 1 site CGATCG
 GGTAGC

SacII: 1 site CCGCGG
 GCGGCC

XbaI: 1 site TCTAGA
 AGATCT

XhoI: 1 site CTCGAG
 GAGCTC

EcoRV: 2 sites GATATC
 CTATAG

NcoI: 2 sites CCATGG
 GGTACC

NdeI: 2 sites CATATG
 GTATAC

SphI: 2 sites GCATCG
 CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4289 End: 4861

Kanr

Start: 6337 End: 6959 (Complementary)

Misc_feature (2 signals)

CMV enhancer

Start: 248 End: 885

SGP(Z)

Start: 1870 End: 4288

Annotations

09000766 054401
10450 69700650

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
ACGCGCGCAA GCGACTACTG CCACTTTGG AGACTGTGTA CGTCGAGGGC
.....
31 GACACGGTCA CAGCTTGCTC GTAAGCGGAT GCCGGGACCA GACAAGCCCG
CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTCCGGC
.....
101 TCAGGCGCGG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
ACTCCCGCGC AGTCGCCCCAC AACCGCCCCAC ACCCCCGACC GAATTGATAC
.....

NdeI

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAATA
CCCGTACTCT CGTCTAACAT GACTCTCAGG TGGTATACCG CACACTTAT
.....
201 CCGCACAGAT CGGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGCCCA
GGCGTGCTA CGCATTCTC TTTTATGCG TAGTCTAACG GATAACCGGT
.....
251 TTGCATACGT TGTATCCATA TCATAATATG TACATTATA TTGGCTCATG
AACGTATGCA ACATAGGTAT ACTATTATAC ATGTAAATAT AACCGAGTAC
.....
301 TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
AGGTTGTAAT GCGCGTACAA CTGTAACATA TAACTGATCA ATAATTATCA
.....
351 AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGCCGCAA
.....
401 ACATAACTTA CCGTAATGCG CCCGCGTGGC TGACCGCCCA ACGACCCCG
TGTATTGAAT GCCATTACCG GCGCGGACCG ACTGGCGGGT TGCTGGGGGC
.....
451 CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
GGGTAACTCG AGTTATTACT GCATACAAGG GTATCATTCG GGTATCCCT
.....
501 CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTG ACGGGTGARC
.....

NdeI

551 GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
CGTCATGTAG TTCACATAGT ATACCGTTCA TCGGGGGGAT AACTCCAGTT
.....
601 TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
ACTGCCATTT ACCGGGGCGA CCGTAATACG GGTCAATGAC TCGAATACCC
.....

NcoI

651 ACTTTCCTAC TTGGCACTAC ATCTACGTAT TAGTCATCGC TATTACCATG
TCAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC
.....

NcoI

701 GTGATCCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTCGACTC
CACTACGCCA AAACCGTCAT GTAGTTAACC GCACCTATCG CCAAACTGAG
.....
751 ACCGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTITT
TGCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAAACAAA
.....
801 GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAAACA CTCCGCCCA
CCGTGGTTTT AGTTGCCCTG AAAGGTTTAA CAGCATTGTT GAGCGGGGGT
.....

851 TTCACGCAAA TGGGCCCTAG GCGGTACGG TGGGAGGTCT ATATAAGCAG
AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTGCTC

901 AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
TCGAGCAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCCACAA

SacII

951 TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCCGAGGC GCCGSCCCTT

1001 CGGTGCATTG GAAACCGCAT TCCCCGTGCC AAGAGTGACC TAAGTACCGC
GCCACGTAAC CTTGCCGCTA AGGGGCACGG TTCTCACTGC ATTCAATGGC

SphI

1051 CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101 TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
AAAACCGAAC CCCGCAATATG TGGGGCGGAA GGAAACGAT ATCCACTACC

1151 TATAGCTTAG CCTATAGGTG TGGGTTATTC ACCATTATTG ACCACTCCCC
ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

1201 TATTGGTGAC GATACTTTC ATTACTAATC CATAACATGG CTCTTTGCCA
ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251 CAACATATCTC TATTGGCTAT ATGCCAATAC TCTGTCTTC AGAGACTGAC
GTTGATAGAG ATAACCGATA TACGGTTATC AGACAGGAAG TCTCTGACTG

1301 ACGGACTCTG TATTTTACA GGATGGGGTC CCATTATTA TTTACAAATT
TGCTGAGAC ATAAAAATGT CCTACCCACG GTAAATAAT AAATGTTAA

1351 CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAACATA
GTGTATATGT TGTTCGGCCA GGGGGCACGG GCGTCAAAA TAATTTGTAT

1401 GCGTGSGATC TCCACGGGAA TCTCGGGTAC GTCTTCGGGA CATGGGCTCT
CGCACCCCTAG AGGTGCGCTT AGACCCCATC CACAAGGCCT GTACCCGAGA

1451 TCTCCGGTAG CGCGGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCCTCC
AGAGGCCATC CCCGCCCTGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

1501 AGCGGCTCAT GGTGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

1551 ACTTAGGCAC AGCACATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
TGAATCCGTG TCGTGTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC

1601 TGGCGGTAGG GTATGTGTCT GAAATGAGC GTGGAGATTG GCTCGGCACG
ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGAGCGTGC

1651 GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATCCAGGCAG
CGACTGCCGC TACCTCTGA ATTCCGTCCG CGTCTTCTTC TACGTCCGTC

1701 CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACCTCC GTTGGGGTCC
GACTCAACAA CATAAGACTA TTCTCACTCT CCATTGAGGG CAACGCCACG

HpaI

1751 TGTAAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCCGG
ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NcoI

1801 CCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
CGCGGGTGGT CTGTATTATC GACTGTCTGA TTCTCTGACA AGGAAAGCTA

NcoIPmlIEcoRVNotI

1851 GGGTCTTTTC TGCAGTCACC GTCOTCGACA CGTGTGATCA CATATCGCGG
CCCAGAAAAG ACCTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCCCG

NotI XbaI

1901 CCGCTCTAGA CCAGGCGCCCT GCATCGAATT GATGAAGATT AAGCCGACAG
GCCGAGATCT GGTCCGCGGA CCTAGCTTAA CTACTTCTAA TTCGGCTGTC

1951 TGAGCGTAAT CTTTCATCTCT CTTAGATTAT TTGTTTCCA GAGTAGGGGT
ACTCGCATTAA GAAGTAGAGA GAATCTAATA AACAAAAGGT CTCATCCCCA

2001 CCTCAGGTCC TTTCGAATCG TGTAAACAAA ATAAACTCCA CTAGAAGGAT
CCAGTCCAGG AAAAGTTAGC ACATTGGTTT TATTGAGGT GATCTTCTA

2051 ATTGTGGGCG AACAAACCAA TGGGCGTTAC AGGAATATTG CAGTTACCTC
TAACACCCCG TTGTTGTGTT ACCCGCAATG TCCTTATAAC GTCAATGGAG

2101 GTGATCGATT CAAGAGGACA TCATTCTTTC TTTGGGTAAT TATCCTTTTC
CACTAGCTAA GTCTCCTGT AGTAAGAAAG AAACCCATTA ATAGGAAAAG

2151 CAAAGAACAT TTCCCATCCC ACTTGGAGTC ATCCACAATA GCACATTACA
GTCTCTGTA AAAGGTAGGG TGAACCTCAG TAGGTCTTAT CGTGAATGT

2201 GGTTAGTGAT GTCGACAAAC TAGTTGTGCG TGACAACTG TCATCCACAA
CCAATCATA CAGCTGTTTG ATCAACAGC ACTGTTTGAC AGTAGGTGTT

2251 ATCAATTGAG ATCAGTTGGA CTGATCTCG AAGGGAATGG AGTGGCAACT
TAGTTAATC TAGTCAACCT GACTTAGAGC TTCCCTTACC TCACCGTGA

2301 GACGTGCCAT CTGCAACTAA AAGATGGGGC TTCAGGTCCG GTGTCCACC
CTGCACGGTA GACGTTGATT TTCTACCCCG AAGTCCAGGC CACAGGGTGG

2351 AAAGGTGGTC AATTATGAAG CTGGTGAATG GGCTCAAAAC TCCTACAATC
TTTCACCAG TTAATACTTC GACCACTTAC CCGACTTTTC ACGATGTTAG

2401 TTGAATCAA AAAACCTGAC GCGAGTGAGT GTCTACCAGC AGCGCCAGAC
AAGTTAGTT TTTTGGACTG CCTCACTCA CAGATGGTGG TCGCGGTCTG

2451 CCGAATCGGG GCTTCCCCCG GTGCCGGTAT GTGCACAAAG TATCAGGAAC
CCCTAAGCCC CGAAGGGCGC CACGGGCATA CACGTGTTTC ATAGTCTTG

2501 GCGACCGTGT GCGGGAGACT TTGCCTTCCA TAAAGAGGGT GCTTTCTTCC
CCCTGCCACA CCGCCTCTGA AACGGAAGGT ATTTCTCCCA CGAAAGAGG

2551 TGTATGATCG ACTTGCTTCC ACAGTTATCT ACCGAGGAAC GACTTTCCGT
ACATACTAGC TGAACGAAGG TGTCAATAGA TGGCTCCTTG CTGAAAGCGA

2601 GAAGGTCTCG TTGCATTCT GATACTGCC CAAGCTAAGA AGGACTTCTT
CTTCACACG AACGTAAAGA CTATGACGG GTTCGATTCT TCCTGAAGAA

2651 CAGCTCACAC CCCTTGAGAG AGCCGGTCAA TGCAACGGAG GACCCCTCTA
GTCCAGTGTG GGGAACTCTC TCGGCCAGTT ACCTTGCTC CTGGGCAGAT

EcoRV

2701 GTGGCTACTA TTCTACCACA ATTAGATATC AGGCTACGG TTTTGAACC
CACCATGAT AAGATGGTGT TAATCTATAG TCCGATGGCC AAAACCTTGG

2751 AATGACACAG ACTACTTCTT CGAGGTTGAC AATTTCACCT ACCTCCAAT
TTACTCTGTC TCATCAACAA GCTCCAATG TTAAACTGGA TGCAGCTTGA

2801 TGAATCAAGA TTCACACCAC AGTTTCTGCT CCAGCTGAAT GAGACAATAT
ACTTAGTCT AAGTGTGGTG TCAAAGACGA GGTGCACTTA CTCTGTTATA

2851 ATACAAGTGG GAAAGGAGC AATACCACGG GAAACTAAT TTGGAAGTTC
TATGTTACCC CTTTCTCTCG TTATGGTGCC CTITTGATTA AACCTTCCAG

2901 AACCCCGAAA TTGATACAAC AATCGGGGAG TGGGCCTTCT GGGAACTAA
TTGGGCTTTT AACTATGTTG TTAGCCCTC ACCCGGAAGA CCCTTGATT

2951 AAAAACCCTA CTAGAAAAAT TCGCAGTGAA GAGTGTCTT TCACAGTTGT
TTTTTGAGT GATCTTTTAA AGCGTCACTT CTCAACAGAA AGTGTCAACA

3001 ATCAAAACGA GCCAAAAACA TCAGTGGTCA GACTCCGGCG CGAAGTTCTT
TAGTTTGCTT CCCTTTTCTG AGTCACCACT CTCAGGCCGC GCTGAAGAA

3051 CCCACCCAGG GACCAACACA ACAACTGAAG ACCACAAAAT CATGCCTTCA
GCCTGGGTCC CTGGTTGTGT TGTGACTTC TGGTGTTTTA GTACCGAAGT

3101 GAAAATTCCT CTGCAATGGT TCAAGTGCAC AGTCAAGGAA GGGAAGCTGC
CTTTAAGGA GACGTTACCA AGTTCACGTG TCAGTTCTCT CCCTTCGAGC

3151 AGTGTCCCAI CTAACAACCC TTGCCACAAT CTCCACGAGT CCCCAATCCC
TCACAGCGTA GATCTTTGGG AACGGTGTAA GAGGTGCTCA GGGGTTAGGG

3201 TCACAACCAA ACCAGGTCCG GACAACAGCA CCCATAATAC ACCGTGTAT
AGTGTGGGT TGGTCCAGGC CTGTTGTGCT GGGTATTATG TGGGCACATA

3251 AAACCTTGACA TCTCTGAGGC AACTCAAGTT GAACAACATC ACCCCAGAAC
TTTGAAGTGT AGAGACTCCG TTGAGTTCAA CTTGTTGTAG TGGCGTCTTG

3301 AGACAACCAAC AGCAGAGCCT CCGACACTCC CTCTGCCACG ACCGCAGCCG
TCTGTTGCTG TCGTGTCCGA GCTGTGAGG GAGACGGTGC TGGCGTCGGC

3351 GACCCCAAAA AGCAGAGAAC ACCAACACGA CCAAGAGCAC TCACTTCCTC
CTGGGGGTTT TCGTCTCTTG TGGTGTGCT CGTTCTCGTG ACTGAAGGAC

3401 GACCCCGCCA CCACAACAAG TCCCCAAAAC CACAGCGAGA CCGCTGGCAA
CTGGGCGGGT CCGTGTGTTT AGCGGTTTTC GTGTGCTCT GCGGACCGTT

3451 CAACAACACT CATCACCAG ATACCGGAGA AGAGAGTGGC AGCAGCCGGA
GTTGTGTGA GTAGTGGTTC TATGGCCTCT TCTCTCAGGC TCGTCCGCTT

3501 AGCTAGGCTT AATTACCAAT ACTATTGCTG GAGTGGCAGG ACTGATCACA
 TCGATCCGAA TTAAAGCTTA TGATAACGAC CTCAGCGTCC TGACTAGTGT

 3551 GCGCGGAGAA GAAGTCGAAG AGAAGCAATT CTCAAATGCTC AACCCAAATG
 CCGCCCTCTT CTTGAGCTTC TCTTCGTTAA CAGTTACGAG TTGGGTTTAC

 3601 CAACCTAAT TTACATTACT GGACTACTCA GGATGAAGGT GCTCCAATCG
 GTTGGGATTA AATGTAATGA CCGATGAGT CCTACTTCCA CGACGTTAGC

 3651 GACTGGCCTG GATACCATAT TTCGGGGCAG CAGCCGAGGG AATTACATA
 CTGACCGGAC CTATGGTATA AAGCCCGGTC GTCGGCTCCC TTAAATGTAT

 3701 GAGCGGCTAA TGCACAATCA AGATGGTTTA ATCTGTGGGT TGAGACAGCT
 CTCGCCGATT ACGTGTTACT TGTACCAAT TAGACACCCA ACTCTGTGCA

 3751 GCGCAACGAG ACCACTCAAG CTCCTCAACT GTTCCTGAGA GCCACAATCG
 CCGGTTGCTC TCGTAGCTTC GAGAAGTTGA CAAGGACTCT CGGTGTTGAC

 3801 ACCTACGCAC CTTTTCATC CTCACCGTA AGGCAATTGA TTTCTGCTG
 TCGATCCGTG GAAAGTTAG GAGTTGGCAT TCCGTTAACT AAAGAACCAC

 3851 CAGCGATGGG GCGGCACATG CCACATTCTG GGACCGGACT CCGTATCGA
 GTCGCTACCC CGCCGTGTAC GGTGTAAGAC CCTGGCCTGA CGACATAGCT

 3901 ACCACATGAT TGGACCAAGA ACATAACAGA CAAAATTGAT CAGATTATTC
 TGCTCTACTA ACCTGGCTCT TGTATTGTCT GTTTTAACTA CTCTAATAAG

 3951 ATGATTTTGT TGATAAAACC CTTCCGGACC AGGGGGACAA TGACAATTGG
 TACTAAACA ACTATTTTGG GAAGGCCTGG TCCCCCTGTT ACTGTTAACC

 4001 TGGACAGGAT GGAGACAATG GATACCGGCA GGTATTGGAG TTACAGGCGT
 ACCTGTCTTA CCTCTGTTAC CTATGCCCGT CCATAACCTC AATGTCCGCA

 4051 TATAATTGCA GTTATCGCTT TATTCTGTAT ATGCAAAATT GTCTTTTAGT
 ATATTAACTG CAATAGCGAA ATAAGACATA TACGTTTAAA CAGAAAATCA

 4101 TTTCTTCAG ATTGCTTCAT GGAAAAGCTC AGCCTCAAT CAATGAAACC
 AAAAGAAGTC TAACGAAGTA CCTTTTCGAG TCGGAGTTTA GTTACTTTGG

 4151 AGGATTTAAT TATATGGATT ACTTGAATCT AAGATTACTT GACAAATGAT
 TCTAAATTA ATATACCTAA TGAACCTAGA TTCTAATGAA CTGTTTACTA

 4201 AATATAATAC ACTGGAGCTT TAAACATAGC CAATGTGATT CTAACCTCCT
 TTATATTATG TGACCTCGAA ATTTGTATCG GTTACACTAA GATTGAGGAA

 4251 TAAACTCACA GTTAATCATA AACAAAGTTT CGAATTGATC TGCTGTGCTT
 ATTTGAGTGT CAATTACTAT TTGTTCCAAA CCTTAAGTAG ACGACACGGA

 4301 TCTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCCGTGC CTTCCTTGAC
 AGATCAACCG TCGGTAGACA ACAAAACGGG AGGGGGCACC GAAGGAACGT

 4351 CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAAATAAAT CAGGAAATG
 CGACCTTCCA CGGTGAGCGT GACAGGAAG GATTATTTTA CTCCTTTAAC

 4401 CATCGCATTG TCTGAGTAGG TGTCAATTCTA TTCTGGGGGG TGGCGTGGGG
 GTACCGTAAC AGACTCATCC ACAGTAAGAT AAGACCCCCC ACCCCACCCC

SphI

4451 CACCACAGCA AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA
GTCGTGTCGT TCCOCTCTCT AACCTTCTG TTATCGTCCG TACGACCCCT

KpnI

4501 TCCGGTGGGC TCTATGGGTA CCCAGGTGCT GAAGAATTGA CCCGGTTCCT
ACGCCACCCG AGATACCCAT GGGTCCACGA CTCCTTAACCT GGGCCAAGGA

4551 CCTGGGCCAG AAAGAAGCAG GCACATCCCC TTCTCTGTGA CACACCCTGT
CGACCCGGTC TTTCTTCGTC CGTGTAGGGG AAGAGACACT GTGTGGGACA

4601 CCACCCCCCT GGTTCCTTAGT TCCAGCCCCA CTCATAGGAC ACTCATAGCT
GGTCCGGCGA CCAAGAATCA AGGTCCGGGT GAGTATCCTG TGAGTATCGA

4651 CAGGAGGGCT CCGCCTTCAA TCCCACCCGC TAAAGTACTT GGAGCGGTCT
GTCTCCCGA GCGGGAAGTT AGGTGGGGC ATTTCATGAA CCTCGCCAGA

4701 CTCCTCCCT CATCAGCCCA CCAAAACCAA CCTAGCCTCC AAGAGTGGGA
GAGCCAGGGA GTAGTCGGGT GGTTCGGTTT GGATCGGAGG TTCTCACCT

4751 AGAAATATAA CCAAGATAGG CTATTAAGTC CAGAGGCAGA GAAATGCCT
TCTTTAATTT CGTTCTATCC GATAATTACG GTCTCCCTCT CTTTACGGA

4801 CCAACATGTC AGGAAGTAAT CAGAGAAATC ATAGAATTTC TTCCGCTTCC
GGTGTACAC TCCTTCATTA CTCTCTTAG TATCTTAAAG AAGCCGAAGC

4851 TCCCTCACTG ACTCGCTCGG CTCGGTCTGT CCGCTCGGC GAGCGGTATC
AGCGAGTGAC TGAGCGAGCG GAGCCAGCAA GCCGACCCCG CTCGCCATAG

4901 AGCTCACTCA AAGCGCGTAA TACGGTTATC CACAGAATCA GGGGATAACG
TCGAGTCAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT CCCCTATTGC

4951 CAGGAAGAA CATGTGAGCA AAAGGCCAGC AAAAGCCAG GAACCGTAA
GTCTTTCTT GTACACTCGT TTTCGGGTG TTTCGGGTG CTTGGCAATT

5001 AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCGCCCCCG CTCAGGAGCA
TTCCGGCGCA ACGACCCCAA AAAGGTATCC GAGCCGGGGG GACTGCTCGT

5051 TCACAAAAAT CGACGCTCAA GTCAGAGCTG GCGAAACCCG ACAGGACTAT
AGTCTTTTAA GCTGCCAGTT CAGTCTCCAC CGCTTTCGGC TGTCTGATA

5101 AAACATACCA GGCCTTTCCG CTTGGAAGCT CCTTCGTGCG CTCTCTGTT
TTTCTATGCT CCCCAGAGG GACCTTCCA GGGAGCACGC CAGAGGACAA

5151 CCGACCCCTCC CGCTTACCGG ATACCTCTCC GCCTTTCTCC CTTGGGGAAG
GGCTCGGACG CCGAATGCCG TATGGACAGG CGGAAGAGG GAAGCCCTTC

5201 CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
GCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA AGCCACATCC

5251 TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC
AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA AGTCCGGCTG

5301 CCGTGGCCTT TATCCGGTAA CTATCGTCTT GATCCAAAC CGGTAAGACA
CGGACGCGGA ATAGGCCATT GATACGAGAA CTCAGGTTGG GCCATTCTGT

```

5351 CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA
    GCTGAATAGC GGTGACCGTC GTCCGGTGACC ATTGTCTTAA TCGTCTCGCT
.....
5401 GGTATGTAGG CCGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACTACGGC
    CCAATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG ATTGATGCCG
.....
5451 TACACTAGAA GCACAGTATT TGGTATCTGC GCTCTGCTGA AGCCACTTAC
    ATGTGATCTT CCTGTCTATA ACCATAGACC CGAGACGACT TCGCTCAATG
.....
5501 CTTCGGAAAA AGAGTTGGTA CCTCTTGATC CGGCAAAACA ACCACCCGTC
    GAAGCCTTTT TCTCAACCAT CGAGAACTAG GCGGTTTGTI TGGTGGCGAC
.....
5551 GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGGC CAGAAAAAAA
    CATCGCCACC AAAAAACAA ACCTTCGTGC TCTAATGCCG GTCTTTTTTT
.....
5601 GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG ACGCTCAGTG
    CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC TCGGAGTCAC
.....
5651 GAACGAAAAA TCACGTTAAG GGATTTTGGT CATGAGATTA TCAAAAAGGA
    CTTGCTTTTG AGTGAATTC CCTAAAACCA GTACTCTAAT AGTTTTCTCT
.....
5701 TCTTCACCTA GATCCTTTTA AATTAATAAT GAAGTTTTAA ATCAATCTAA
    AGAAGTGGAT CTAGAAAAAT TTAATTTTTA CTCAAAAT TAGTTAGATT
.....
5751 AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA
    TCATATATAC TCATTTGAAC CAGACTGTCA ATGATTACGA ATTAGTCACT
.....
5801 GCCACCTATC TCAGCGATCT GTCTAATTCG TTCATCCATA GTTGCTGAC
    CCGTGGATAG AGTCGCTAGA CAGATAAAGC AAGTAGGTAT CAACGGACTG
.....
5851 TCCCGGGGGG GGGGGCGCTG AGGTCTGCCT CGTGAAGAAG GTCTTGCTGA
    ACGCCCCCCC CCCCCCGGAC TCCAGACGGA GCACCTCTTC CACACGACT
.....
5901 CTCATACCAG CCGTGAATCG CCCCATCATC CAGCCAGAAA GTGAGGGAGC
    GAGTATGGTC CGGACTTAGC GGGGTAGTAG GTCGGTCTTT CACTCCCTCG
.....
5951 CACCGTTGAT GAGAGCTTTG TTGTAGGTGG ACCAGTTGGT GATTTTGAAC
    GTGCCAACTA CTCTCGAAAC AACATCCACC TGCTCAACCA CTAAAACTTG
.....
6001 TTTCGCTTTG CCACGGAACG GTCTGCGTTG TCGGGAAGAT GCGTGATCTG
    AAAACGAAAC GGTGCCCTGC CAGACGCAAC ACCCCTTCTA CGCACTAGAC
.....
6051 ATCCTTCAAC TCAGCAAAAG TTCGATTTAT TCAACAAAGC CGCCGTCCCG
    TAGGAAGTTG AGTCGTTTTT AAGCTAAATA AGTTGTTTCG GCGGCAGGGC
.....
6101 TCAAGTCAGC GTAATGCTCT CCCAGTGTTA CAACCAATTA ACCAATTCG
    AGTTCAGTCG CATTACGAGA CCGTCACAA GTTGCTTAAT TGGTTAAGAC
.....
6151 ATTAGAAAAA CTCATCGAGC ATCAAAATGA ACTGCAATTT ATTCATATCA
    TAATCTTTTT GAGTAGCTCG TAGTTTACTT TGACGTTAAA TAAGTATAGT
.....
6201 GGATTATCAA TACCATATTT TTGAAAAAGC CGTTTCTGTA ATGAAGGAGA
    CCTAATAGTT ATGGTATAAA AACTTTTTTCG GCAAAGACAT TACTTCCTCT
.....
6251 AACTCACCG AGGCAGTTCC ATAGGATGCC AAGATCCTGG TATCGGTCTG
    TTTGAGTGGC TCCGTCAAGG TATCCTACCG TTCTAGGACC ATAGCCAGAC
.....

```

6301 CGATTCCGAC TCGTCCAACA TCAATACAAC CTATTAATTT CCCCTCGTCA
GCTAAGGCTG ACCAGGTGTG AGTTATGTTG GATAATTAAA GGGGAGCAGT

6351 AAAATAAGGT TATCAAGTGA GAAATCACCA TGAGTCACGA CTGAATCCGG
TTTTATTCCA ATAGTTCACT CTTAGTGGT ACTCACTGCT GACTTAGGCC

HindIII

6401 TGAGAATGGC AAAAGCTTAT GCATTTCTTT CCAGACTTGT TCAACAGGCC
ACTCTTACCG TTTTCGAATA CGTAAAGAAA GGTCTGAACA AGTTGTCCGG

6451 AGCCATTACG CTCGTCATCA AAATCACTCG CATCAACCAA ACCGTTATTC
TCGGTAATGC GAGCACTAGT TTAGTGTAGC GTAGTTGGTT TGGCAATAAG

FvuI

6501 ATTCTGTATT GCGCCTGAGC GACACGAAAT ACGCGATCGC TGTAAAAGG
TAAGCACTAA CCGGGACTCG CTCTGCTTTA TCGCCTAGCG ACAATTTTCC

6551 ACAATTACAA ACAGGAATCG AATGCAACCG GCGCAGGAAC ACTGCCAGCG
TGTTAATGTT TGTCTTACG TTACGTTGGC CGCGTCTTG TGACGGTCGC

6601 CATCAACAAT ATTTTCACCT GAATCAGGAT ATTCTTCTAA TACCTGGAAT
GTAGTGTGTA TAAAAGTGGC CTTAGTCTTA TAAGAAGATT ATGGACCTTA

6651 GCTGTTTTCC CGGGGATCCG AGTGGTGAGT AACCATGCAT CATCAGGAGT
CGACAAAAGG GCCCCTAGCG TCACCACTCA TTGGTACGTA GTAGTCTCA

6701 ACGGATAAAA TGCTTCATGG TCGGAAGAGC CATAAATTCC GTCAGCCAGT
TGCTTATTTT ACGAACTACC AGCCTTCTCC GTATTAAAGG CAGTCGGTCA

6751 TTACTCTGAC CATCTCATCT GTAACATCAT TGGCAACGCT ACCTTTGCCA
AATCAGACTG GTACAGTAGA CATGTAGTA ACCGTTGCGA TGGAAACGGT

6801 TGTTTCAGAA ACAACTCTGG CGCATCGGGC TTCCCATACA ATCGATAGAT
ACAAAGTCTT TGTGTAGACC GCGTAGCCCG AAGGGTATGT TAGCTATCTA

6851 TGTCCACCT GATTGCCCGA CATTATCGCG AGCCCATTTA TACCCATATA
ACAGCGTGGC CTAACGGGCT GTAATAGCGC TCGGGTAAAT ATGGGTATAT

XhoI

6901 AATCAGCATC CATCTTGGAA TTTAATCGCG GCCTCGAGCA AGACGTTTCC
TTAGTCGTAG GTACAACCTT AAATTAGCGC CGGAGCTCGT TCTGCAAAGG

6951 CGTTGAATAT GGCTCATAAC ACCCCTTGTA TTAAGTTTA TGTAAAGAGA
GCAACTTATA CCGAGTATTG TGGGGAACAT AATGACAAAT ACATTCTCT

7001 CAGTTTTATT GTTCATCATG ATATATTTTT ATCTTGTCCA ATGTAAATC
GTCAAAATAA CAAGTACTAC TATATAAAA TAGAACAGT TACATTGTAG

DraIII

7051 AGAGATTTTG AGACACAACG TGGCTTTCCG CCCCCCCCCA TTATTGAAGC
TCTCTAAAAC TCTGTGTTGC ACCGAAAGGG GGGGGGGGGT AATAACTTCG

7101 ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA
TAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT

7151 GAAAAATAA CAATAGGGG TCCGCGCAC ATTCCCGA AAAGTCCAC
CTTTTATTT GTTTATCCC AAGCGCGTG TAAAGGGCT TTTCAGGTG

7201 CTGACGTCTA AGAACCATT ATTATCATGA CATTACCTA TAAAAATAGG
GACTGCAGAT TCTTTGGTAA TAATACTACT GTAATTGGAT ATTTTATCC

7251 CGTATCACGA GGGCCTTTCG TC
GCATAGTGCT CCGGGAAGC AG

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/01382

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : A61K 48/00; C07H 21/04; C12N 15/63, 15/86, 5/10, 15/40

US CL : Please See Extra Sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/450, 93.2, 93.21; 536/23.1, 23.72; 435/5, 6, 455, 457, 458, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog, Biosis, Medline, Biotech

Search terms: Ebola virus, glycoprotein, transmembrane, targeting

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,166,320 A (WU et al.) 24 November 1992, columns 9-10 and claims 1-18.	1-20
A	SANCHEZ et al. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. Proceedings of the National Academy of Sciences. April 1996, Vol. 93, pages 3602-3607, especially page 3604.	3, 12
A	FELGNER et al. Lipofection: A highly efficient, lipid-mediated DNA-transfection procedure. Proceedings of the National Academy of Sciences. November 1987, Vol. 84, pages 7413-7417, especially page 7414.	1-20

☒ Further documents are listed in the continuation of Box C ☐ See patent family annex.

* Special categories of cited documents	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*&* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

12 APRIL 1999

Date of mailing of the international search report

11 MAY 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

DAVID GUZO

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/01382

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	VOLCHKOV et al. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. Virology. 1995, Vol. 214, pages 421-430, especially page 424.	3, 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/01382

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

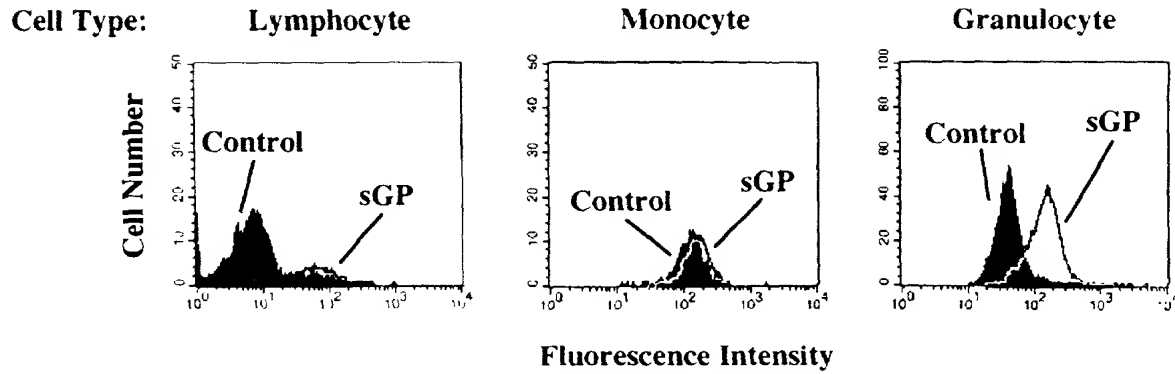
424/450, 93.2, 93.21; 536/23.1, 23.72; 435/5, 6, 455, 457, 458, 320.1

1/11

Figure 1A

Figure 1A1

Figure 1A2



Antibody: Ig
Ebola Protein: Control

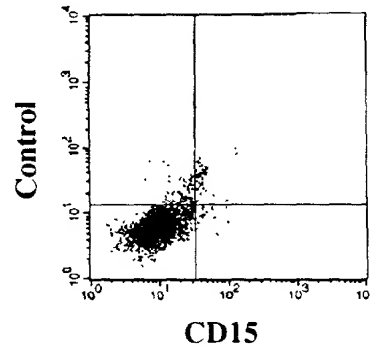


Figure 1B

α -CD15
sGP

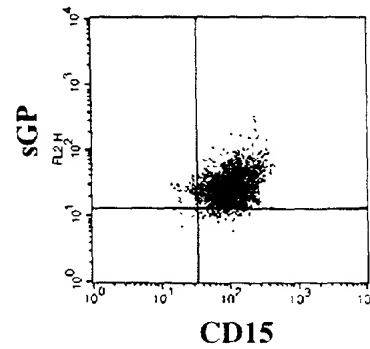


Figure 1B1

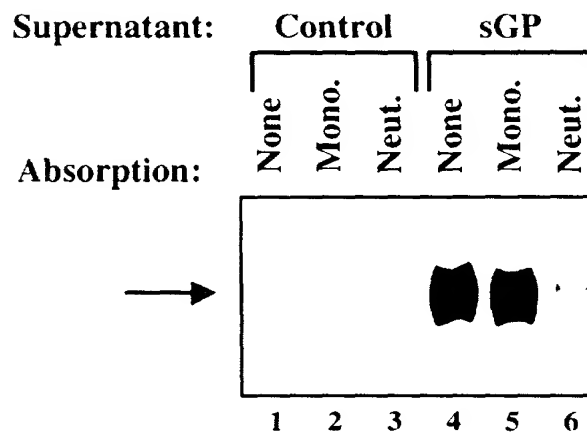


Figure 1C

09/600766

2/11

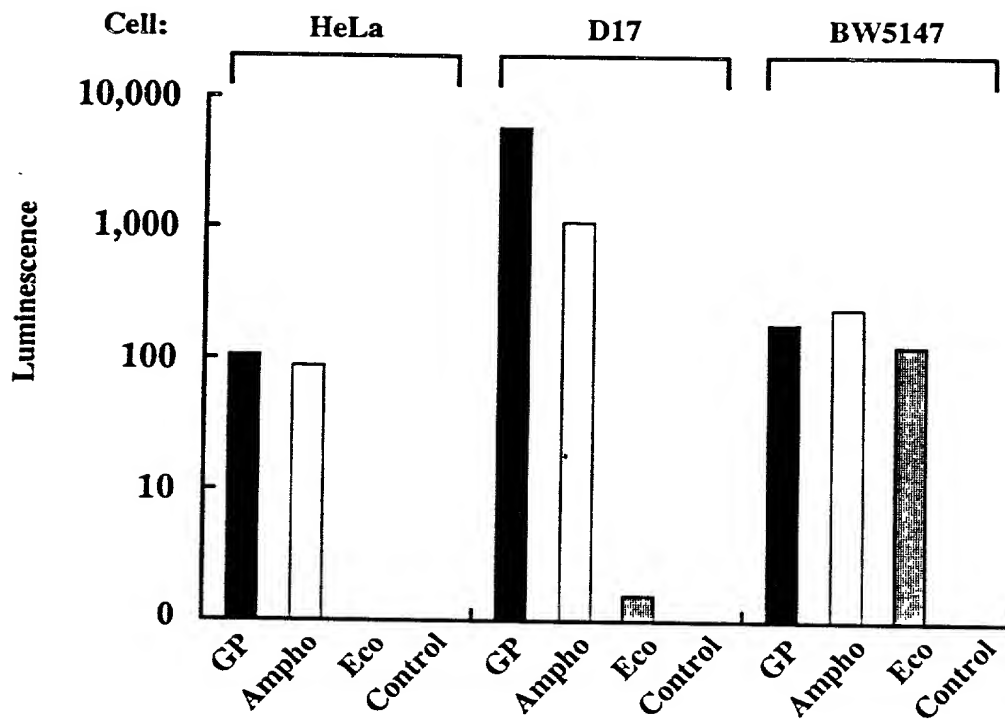


Figure 2A

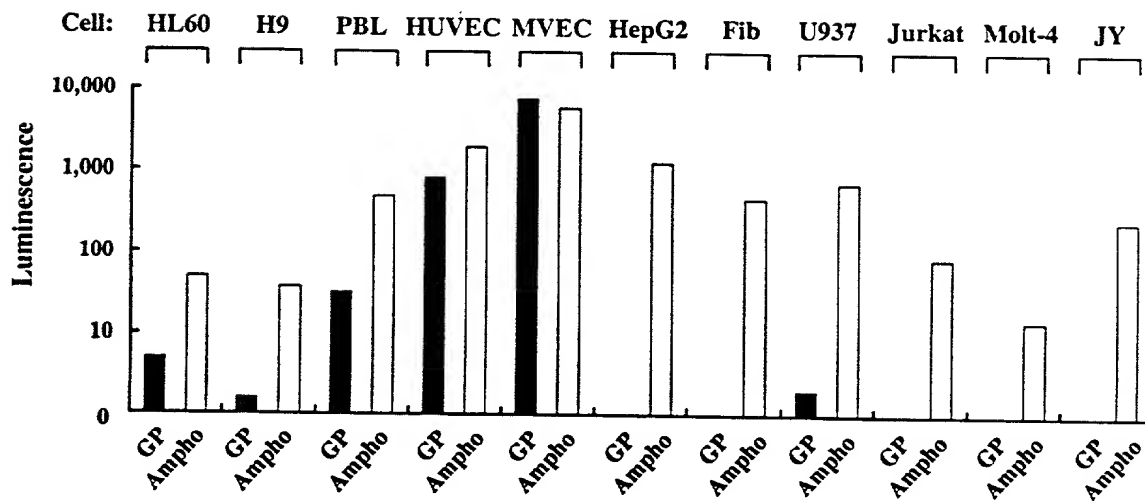
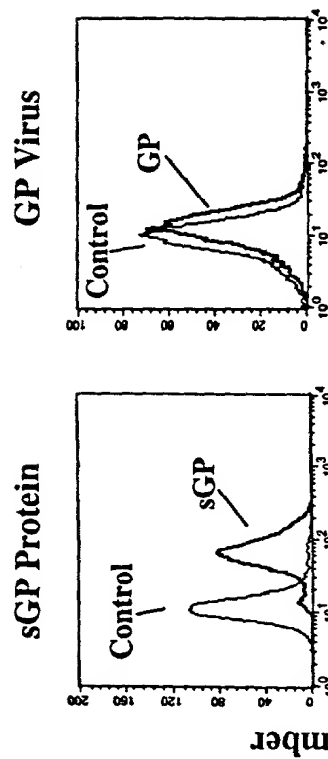


Figure 2B

Figure 2C1



GP Virus

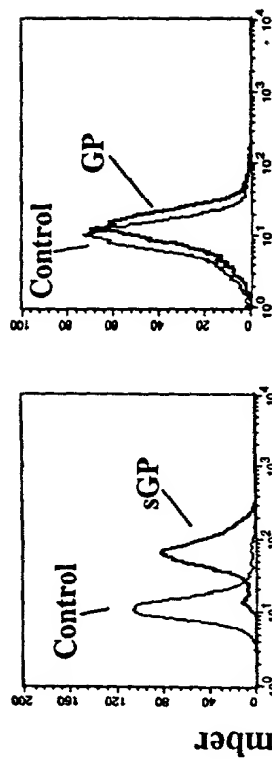
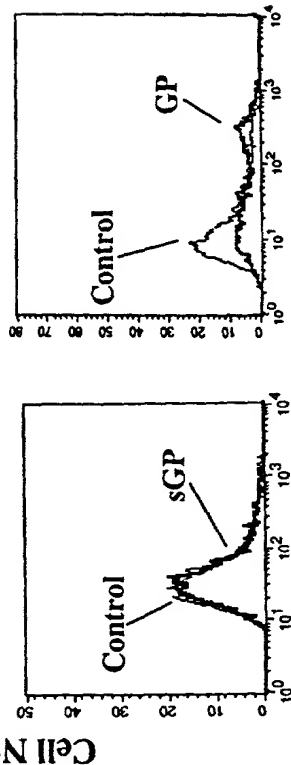


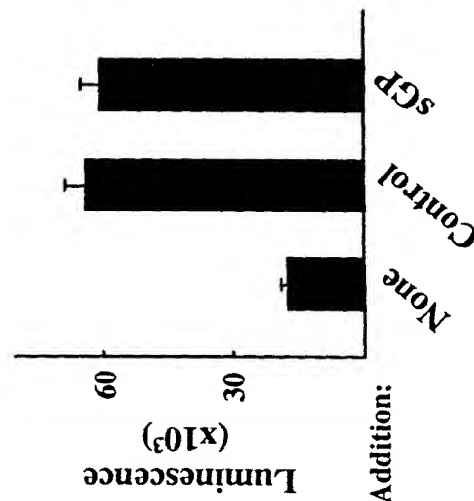
Figure 2C

Figure 2C3



Fluorescence Intensity

Figure 2D



4/11

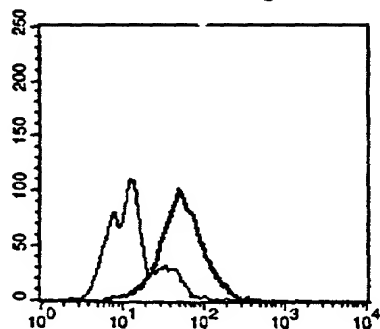
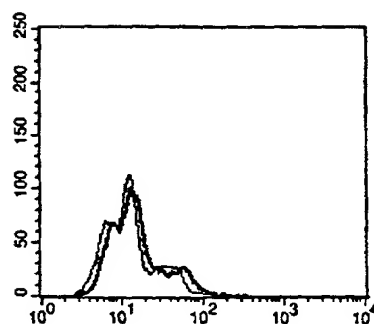
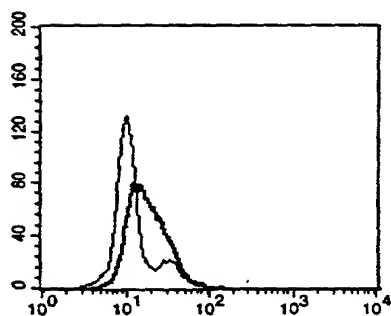
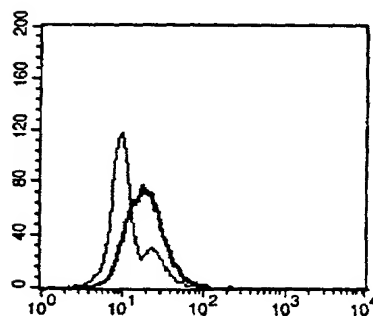
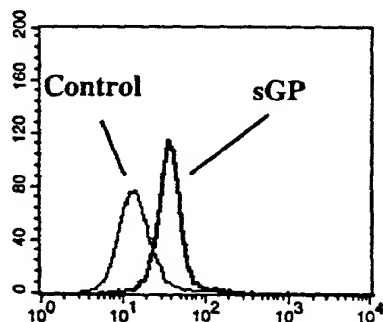
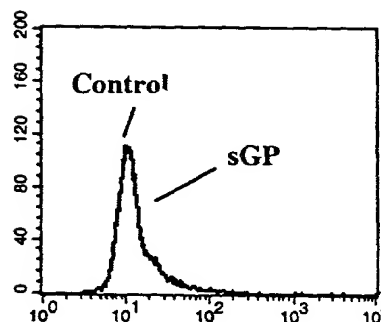
Figure 3A**Ab: Control (Ig)****Figure 3B** **α -CD16****Figure 3C****Ab: Control (Ig)****Cell Number****Figure 3D** **α -CD62L****Figure 3E****Stimulation: None****Figure 3F****PMA****Fluorescence Intensity**

Figure 4A

Pre-stimulation

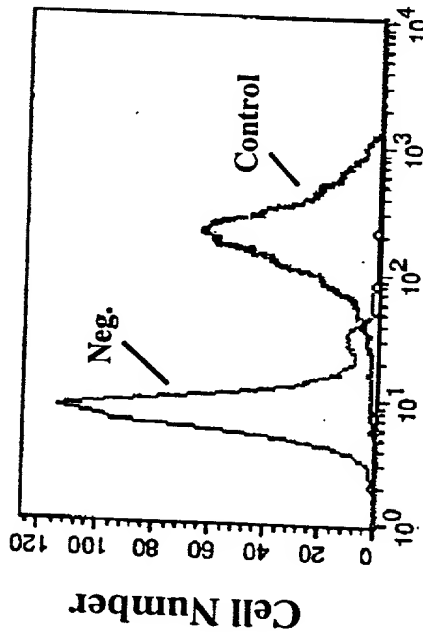
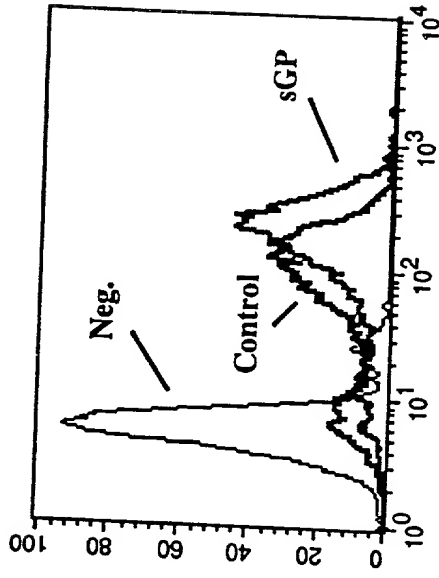


Figure 4B

Post-stimulation



Fluorescence Intensity

Figure 5A

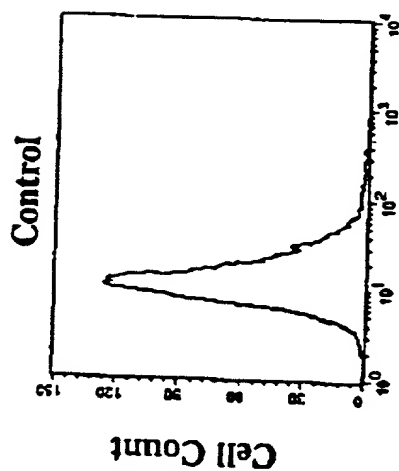


Figure 5B

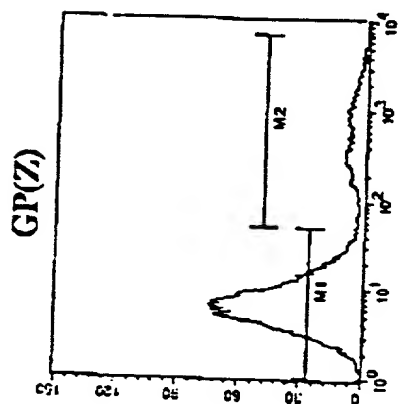
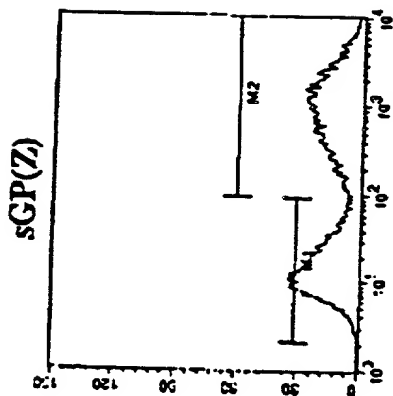


Figure 5C



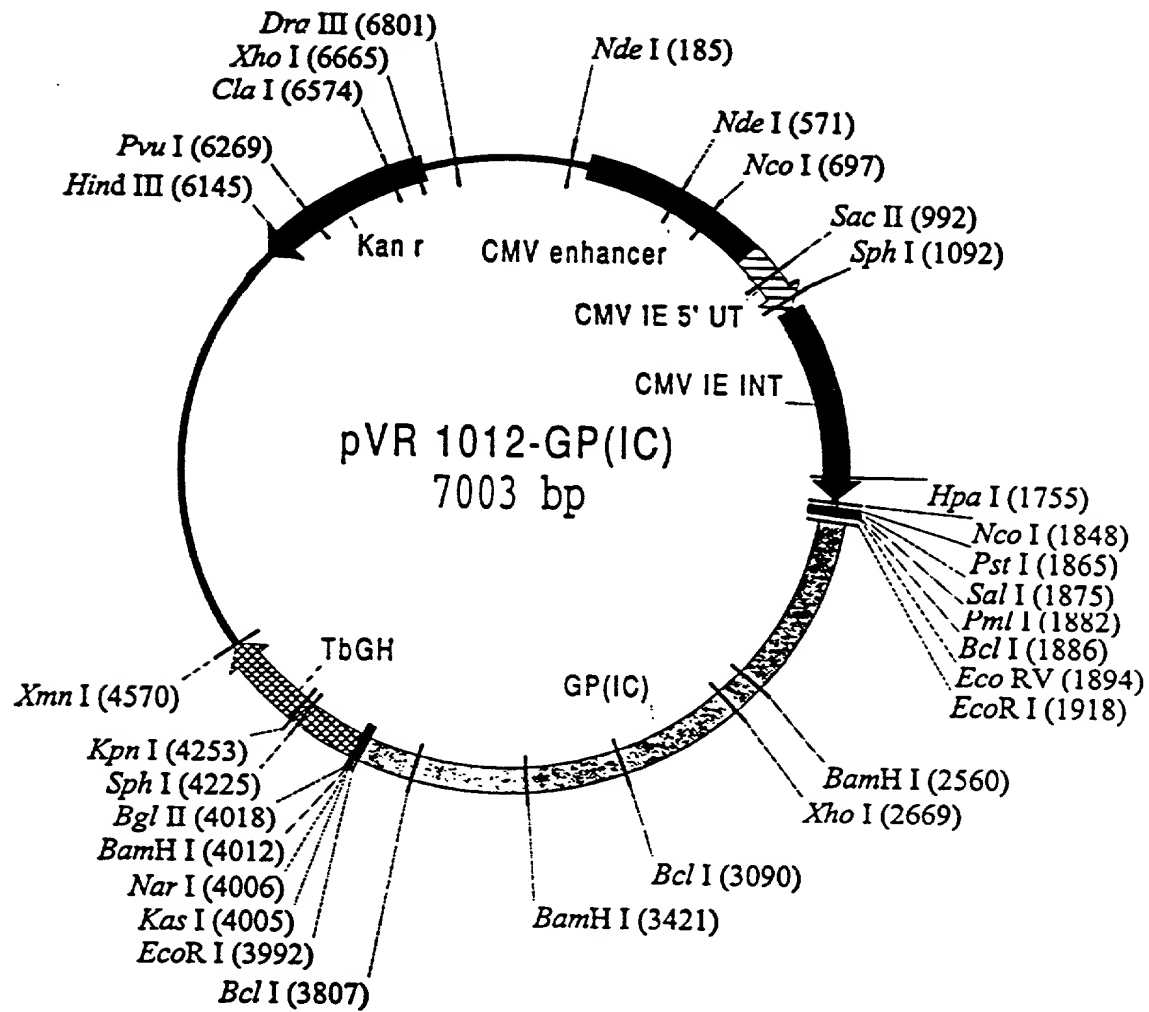


Figure 6

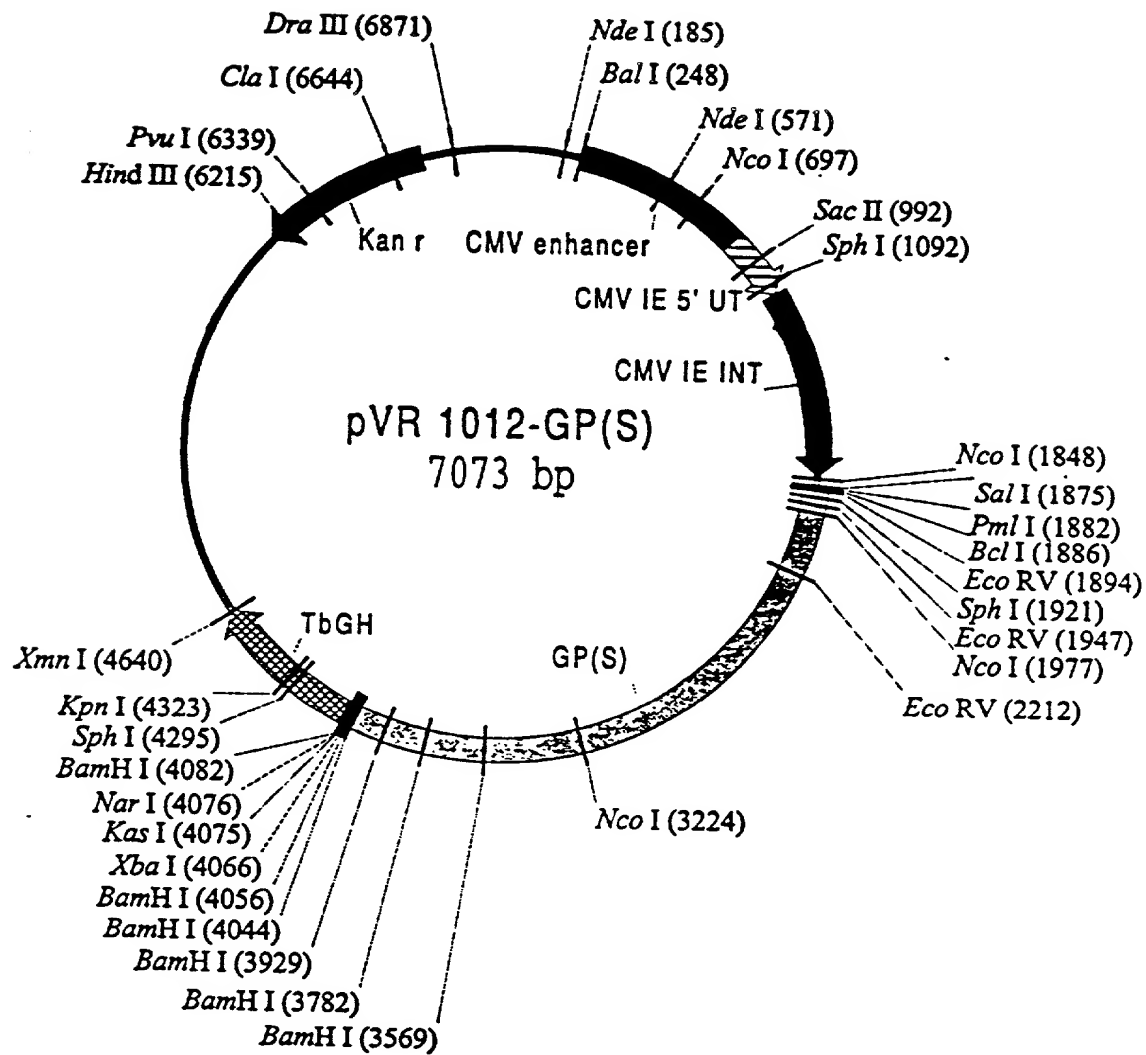
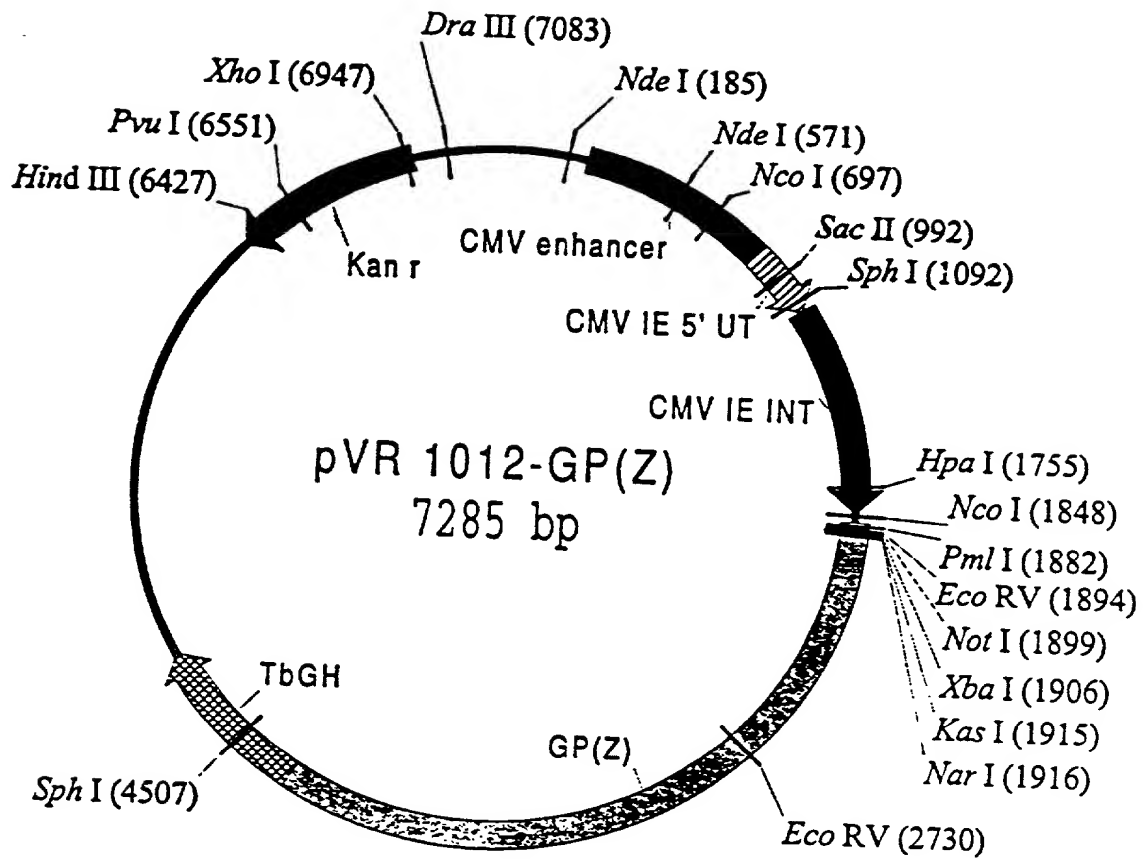


Figure 7

**Figure 8**

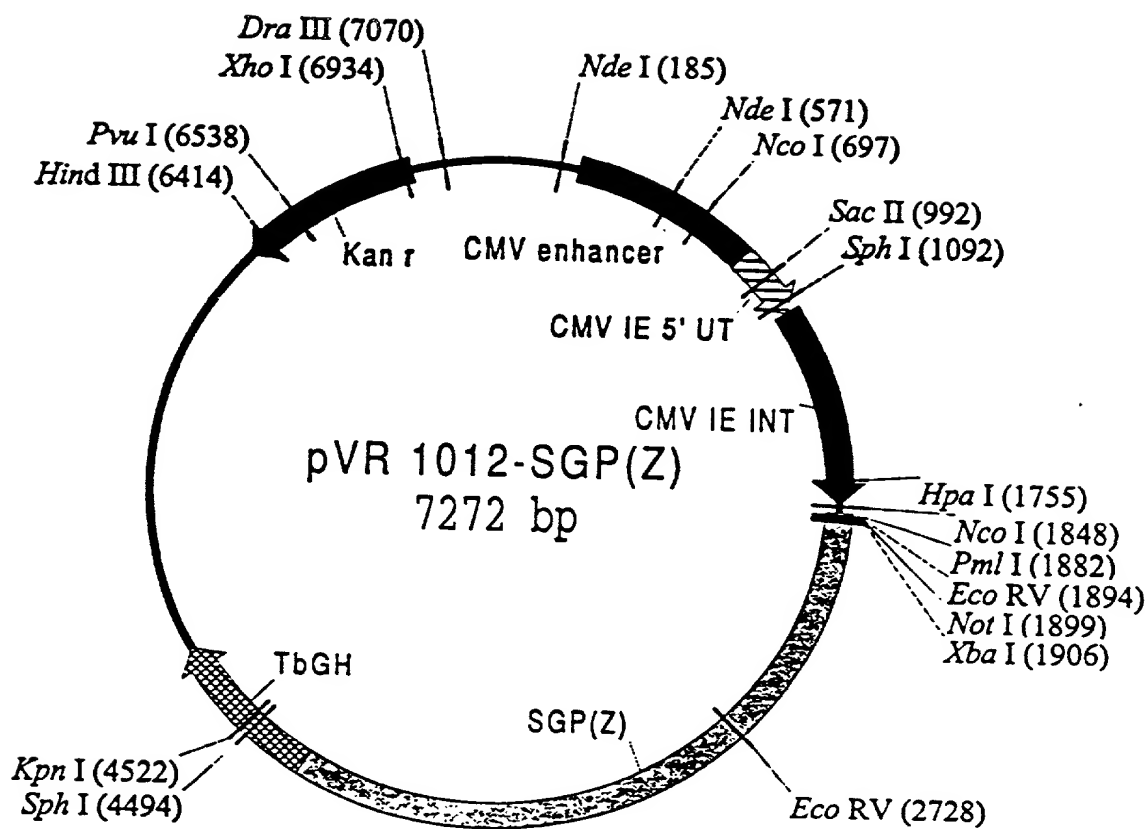


Figure 9

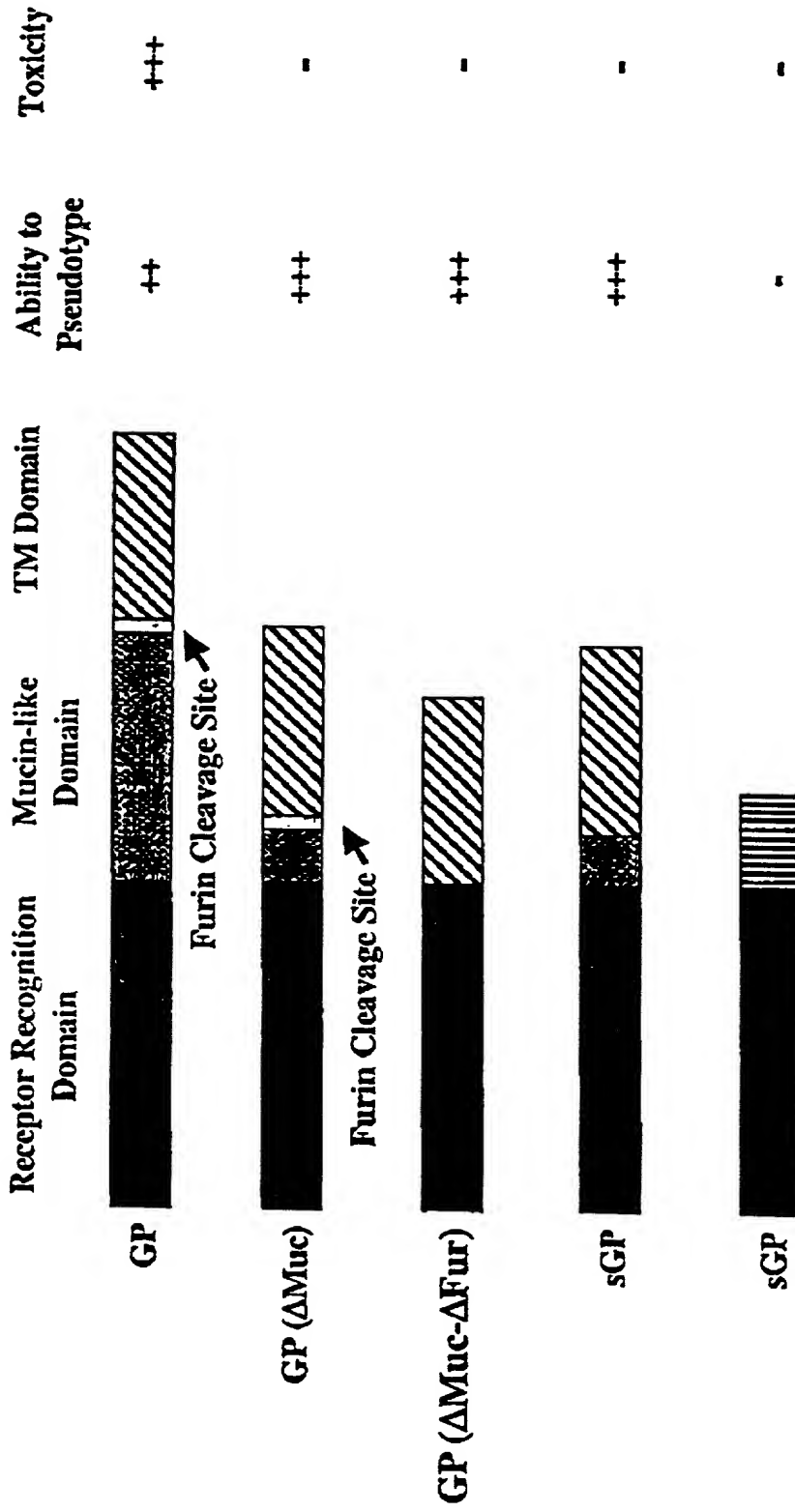


Figure 10



**DECLARATION, PETITION AND POWER OF ATTORNEY
FOR PATENT APPLICATION**

(Check one):

- ☐ Declaration Submitted with Initial Filing
☒ Declaration Submitted after Initial Filing

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

**TARGETING GENE TRANSFER VECTORS TO CERTAIN CELL
TYPES BY PSEUDOTYPING WITH VIRAL GLYCOPROTEIN**

the specification of which (check one):

- ☐ is attached hereto.
OR
☒ was filed on 21 January 1999 as PCT International Application Number
PCT/US99/01382
☐ and was amended by PCT Article 19 Amendment on _____
(if applicable),
☐ and was amended by PCT Article 34 Amendment on _____
(if applicable).

I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby state that I have reviewed and understood the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

09600766-051401

PRIORITY CLAIM

(Check one):

- ☐ no such applications have been filed.
- ☒ such applications have been filed as follows

1) FOREIGN PRIORITY CLAIM: I hereby claim foreign priority benefits under Title 35, United States Code, §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate or §365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (dd/mm/yyyy)	Priority Not Claimed	Certified Copy Attached	
				Yes	No
			<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

☐ Additional foreign application numbers are listed on a supplemental priority sheet attached hereto.

2) PROVISIONAL PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

Provisional Application Number(s)	Filing Date (dd/mm/yyyy)
60/072,033	21 January 1998 (21.01.98)

☐ Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

3) U.S./PCT PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code, §120 of any United States application or §365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose information which is known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application Number	PCT Parent Number	Parent Filing Date (dd/mm/yyyy)	Parent Patent Number (if applicable)
	PCT/US99/01382	21 January 1999 (21.01.99)	

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority

09600765 "051491



Atty Docket No.: UMV-1474US

**DECLARATION, PETITION AND POWER OF ATTORNEY
FOR PATENT APPLICATION**

(Check one):

- ☐ Declaration Submitted with Initial Filing
☒ Declaration Submitted after Initial Filing

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

**TARGETING GENE TRANSFER VECTORS TO CERTAIN CELL
TYPES BY PSEUDOTYPING WITH VIRAL GLYCOPROTEIN**

the specification of which (check one):

- ☐ is attached hereto.
OR
☒ was filed on 21 January 1999 as PCT International Application Number
PCT/US99/01382
☐ and was amended by PCT Article 19 Amendment on _____
(if applicable),
☐ and was amended by PCT Article 34 Amendment on _____
(if applicable).

I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby state that I have reviewed and understood the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

PRIORITY CLAIM

(Check one):

- ☐ no such applications have been filed.
- ☒ such applications have been filed as follows

1) FOREIGN PRIORITY CLAIM: I hereby claim foreign priority benefits under Title 35, United States Code, §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate or §365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (dd/mm/yyyy)	Priority Not Claimed	Certified Copy Attached	
				Yes	No
			<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

- ☐ Additional foreign application numbers are listed on a supplemental priority sheet attached hereto.

2) PROVISIONAL PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

Provisional Application Number(s)	Filing Date (dd/mm/yyyy)
60/072,033	21 January 1998 (21.01.98)

- ☐ Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

3) U.S./PCT PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code, §120 of any United States application or §365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose information which is known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application Number	PCT Parent Number	Parent Filing Date (dd/mm/yyyy)	Parent Patent Number (if applicable)
	PCT/US99/01382	21 January 1999 (21.01.99)	

- ☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority

09600766-051401

POWER OF ATTORNEY:

As a named inventor, I hereby appoint the following attorneys and/or agents to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

James E. Cockfield Reg. No. 19,162
 Thomas V. Smurzynski Reg. No. 24,798
 Ralph A. Loren Reg. No. 29,325
 Giulio A. DeConti, Jr. Reg. No. 31,503
 Ann Lamport Hammitte Reg. No. 34,858
 Elizabeth A. Hanley Reg. No. 33,505
 Amy E. Mandragouras Reg. No. 36,207
 Anthony A. Laurentano Reg. No. 38,220
 Jane E. Remillard Reg. No. 38,872
 Jeremiah Lynch Reg. No. 17,425
 Kevin J. Canning Reg. No. 35,470
 David A. Lane, Jr. Reg. No. 39,261
 Jeanne M. DiGiorgio Reg. No. 41,710

Megan E. Williams Reg. No. 43,270
 Nicholas P. Triano III Reg. No. 36,397
 Peter C. Lauro Reg. No. 32,360
 DeAnn F. Smith Reg. No. 36,683
 William D. DeVaul Reg. No. 42,483
 David J. Ridders Reg. No. 43,882
 Chi Suk Kim Reg. No. 42,728
 Maria Laccotripe Zacharakis Limited Recognition
 Under 37 C.F.R. § 10.9(b)
 Debra J. Milasincic Reg. No. P46,931
 David R. Burns Reg. No. P46,590

Send Correspondence to:

DeAnn F. Smith, Lahive & Cockfield, LLP, 28 State Street, Boston, Massachusetts 02109, United States of America

Direct Telephone Calls to: (name and telephone number)

DeAnn F. Smith, (617) 227-7400

Wherefore I petition that letters patent be granted to me for the invention or discovery described and claimed in the attached specification and claims, and hereby subscribe my name to said specification and claims and to the foregoing declaration, power of attorney, and this petition.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor Gary J. NABEL	
Inventor's signature	Date
Residence 2520 30th Street, N.W., Washington, D.C. 20008	
Citizenship United States of America	
Post Office Address (if different)	

00600766-051401

POWER OF ATTORNEY:

As a named inventor, I hereby appoint the following attorneys and/or agents to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

James E. Cockfield Reg. No. 19,162
 Thomas V. Smurzynski Reg. No. 24,798
 Ralph A. Loren Reg. No. 29,325
 Giulio A. DeConti, Jr. Reg. No. 31,503
 Ann Lamport Hammitte Reg. No. 34,858
 Elizabeth A. Hanley Reg. No. 33,505
 Amy E. Mandragouras Reg. No. 36,207
 Anthony A. Laurentano Reg. No. 38,220
 Jane E. Remillard Reg. No. 38,872
 Jeremiah Lynch Reg. No. 17,425
 Kevin J. Canning Reg. No. 35,470
 David A. Lane, Jr. Reg. No. 39,261
 Jeanne M. DiGiorgio Reg. No. 41,710

Megan E. Williams Reg. No. 43,270
 Nicholas P. Triano III Reg. No. 36,397
 Peter C. Lauro Reg. No. 32,360
 DeAnn F. Smith Reg. No. 36,683
 William D. DeVaul Reg. No. 42,483
 David J. Ridders Reg. No. 43,882
 Chi Suk Kim Reg. No. 42,728
 Maria Laccotripe Zacharakis Limited Recognition
 Under 37 C.F.R. § 10.9(b)
 Debra J. Milasincic Reg. No. P46,931
 David R. Burns Reg. No. P46,590

Send Correspondence to:

DeAnn F. Smith, Lahive & Cockfield, LLP, 28 State Street, Boston, Massachusetts 02109, United States of America

Direct Telephone Calls to: (name and telephone number)

DeAnn F. Smith, (617) 227-7400

Wherefore I petition that letters patent be granted to me for the invention or discovery described and claimed in the attached specification and claims, and hereby subscribe my name to said specification and claims and to the foregoing declaration, power of attorney, and this petition.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor Gary J. NABEL	
Inventor's signature <i>Gary J. Nabel</i>	Date <i>4/25/01</i>
Residence 2520 30th Street, N.W., Washington, D.C. 20008	
Citizenship United States of America	
Post Office Address (if different)	

09600766-051401

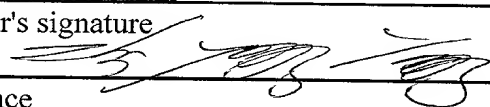
2W

Full name of sole or first inventor Rafael DELGADO	
Inventor's signature <i>R Delgado</i>	Date
Residence PASEO DE SOMOSTERRA 30 TRES CANTOS ; 28761 , MADRID , SPAIN ESX	
Citizenship Spain	
Post Office Address (if different)	

Full name of sole or first inventor Zhi-yong YANG	
Inventor's signature	Date
Residence 10004 Penfold Ct., Potomac, Maryland 20854	
Citizenship China	
Post Office Address (if different)	

00000766 051401

Full name of sole or first inventor Rafael DELGADO	
Inventor's signature	Date
Residence	
Citizenship Spain	
Post Office Address (if different)	

Full name of sole or first inventor Zhi-yong YANG	
Inventor's signature 	Date 3/26/2001
Residence 10004 Penfold Ct., Potomac, Maryland 20854 U.S.	
Citizenship China	
Post Office Address (if different)	

3W
09600766 " 051401
T04750 " 930066